Department of Physics and Astronomy, University of Canterbury,
Private Bag 4800, Christchurch, New Zealand

A Patient Position Guidance System in Radiotherapy Using Augmented Reality

James Talbot (BSc)

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in the University of Canterbury

Supervisor: Dr. Juergen Meyer
Co-supervisors: Dr. Richard Watts, Dr. Raphaël Grasset

Abstract

A system for visual guidance in patient set-up for external-beam radiotherapy procedures was developed using augmented reality. The system uses video cameras to obtain views of the linear accelerator, and the live images are displayed on a monitor in the treatment room. A 3D model of the patient's external surface, obtained from planning CT data, is superimposed onto the treatment couch in the camera images. The augmented monitor can then be viewed, and alignment performed against the virtual contour.

The system provides an intuitive method for set-up guidance, and allows non-rigid deformations to patient pose to be visualised. It also allows changes to patient geometry between treatment fractions to become observable, and can remain in operation throughout the treatment procedure, so that patient motion becomes apparent.

Coordinate registration between the camera view and the linac is performed using a cube which is aligned with the linac isocentre using room lasers or cone-beam CT. The AR tracking software detects planar fiducial tracking markers attached to the cube faces, and determines their positions in order to perform pose estimation of the 3D model on-screen.

Experimental results with an anthropomorphic phantom in a clinical environment have shown that the system can be used to position a rigid-body with a translational error of 3 mm, and a rotational error of 0.19°, 0.06° and 0.27°, corresponding to pitch, roll and yaw respectively.

With further developments to optimise the system accuracy and its interface, it could be made into a valuable tool for radiotherapy clinics. The outcome of the project has been encouraging, and has shown that augmented reality for patient set-up guidance has great potential.

Acknowledgements

First of all, I would like to express immense gratitude towards my supervisors at the University of Canterbury, Dr. Juergen Meyer and Dr. Richard Watts. Their advice and dedication to this project saw that it was steered in the right direction, and ensured that it became substantially more than merely a means for obtaining a qualification. Thanks to their encouragement, I have been able to present this work at conferences, and hope to have the research published in a prominent journal over the coming months.

Equally deserving of thanks are the people of HITLabNZ, in particular Dr. Raphaël Grasset, without whose assistance this project would have been entirely impossible. It is disconcerting to think how early we would have reached a dead end in this work without Raphaël's knowledge and expertise in the field of augmented reality.

Thanks also go to Dr. Mark Bird and Dr. Jerome Gastaldo for their assistance with equipment during the clinical set-up at Christchurch Hospital, and to Graeme Kershaw at the Department of Physics and Astronomy Mechanical Workshop for constructing the "registration cube," an integral component of the system.

Finally, thank you to my parents, Alison and Lindsay, my brothers, Gareth and Oren, and to my various office colleagues and flatmates throughout the year: Stuart, Greg, Latoya, James, Bernadette, Jidi, Mike, Laura and Matty. Their assistance and contributions throughout the duration of the project, however small or large, helped immeasurably.

Contents

Chapter 1

Introduction

1.1 Background

Cancer is caused by the uncontrolled proliferation of cells in the body, and is one of the leading causes of death worldwide [1]. As of 2005, around 16,000 New Zealanders develop cancer each year. 7500 of these cases are fatal [2]. Globally, 7.6 million people died of cancer in 2005, which constituted 13% of worldwide deaths that year [1]. Due to rising population and life expectancy, these figures have been increasing, with an expected 22,000 cases and 9000 deaths in New Zealand in 2011 [2]. Worldwide annual cancer mortality is expected to rise to 9 million in 2015 and to 11.5 million by 2030 [1].

Along with chemotherapy and surgery, radiotherapy has a large role in the treatment and management of cancer. In New Zealand and Australia it is estimated that around 50% of cancer patients require radiation treatment for curative or palliative reasons [3, 4].

Radiotherapy is a capital-intensive treatment modality, and requires highly specialised staff to plan and carry out procedures. With the combined expenses of capital investment and operational costs incurred by a radiotherapy facility, an entire course of treatment can cost as much as NZ$7600 (as of 2001 [3]). Despite this figure, radiation treatment is known to be no less cost-effective than surgery or chemotherapy [3].

Increasing cancer rates have caused the field of radiotherapy to develop significantly over the last decade. With expanding radiotherapy clinics and modern advances in computing,

technology in the field has seen major developments, and research to improve treatment procedures is ongoing [5].

1.2 Radiotherapy

Radiotherapy is the medical implementation of ionizing radiation to cause biological damage to cancer cells. Depending on the nature of the tumour, the radiation dose may be delivered from an external source at a distance from the patient (external beam radiotherapy) or from a radionuclide source placed directly into or close to the volume of the tumour (brachytherapy). This thesis focuses on the technicalities involved in external beam procedures.

In most modern external beam treatments, the radiation is administered in the form of x-rays, produced by a medical linear accelerator (linac). It is a complex process, and requires a dedicated staff of radiation oncologists, therapists and physicists to ensure that radiation is delivered to the patient safely and effectively. There are a number of steps which comprise the entire process, each of which is essential to guarantee the quality of treatment [6]. The steps involved in a typical radiotherapy procedure are described in Figure 1.1.

After a patient has been diagnosed and a radiotherapy course is prescribed, anatomical patient data is obtained, typically from a computed tomography (CT) scan. The data is transferred to a treatment planning station, where an appropriate course of therapy is established.

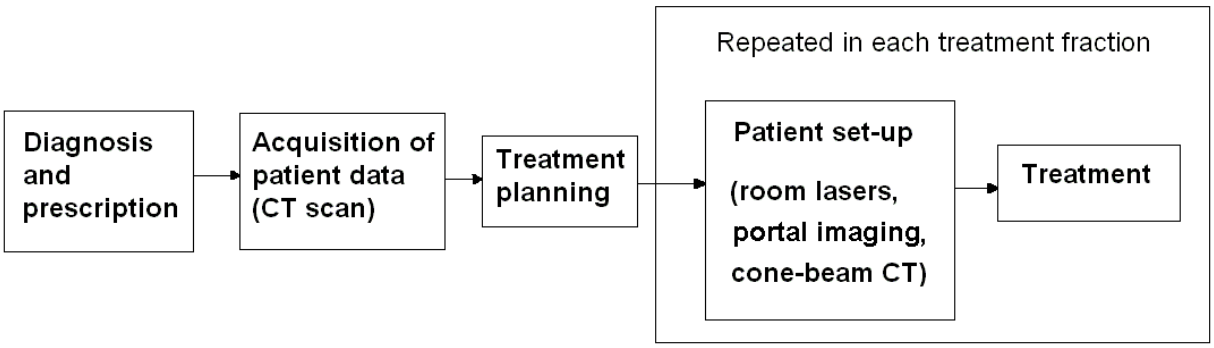


Figure 1.1: Workflow of a typical radiotherapy procedure.

When treatment is carried out on the linear accelerator, the patient must be aligned on the treatment couch in a way that replicates the planning CT. This ensures that the radiation beams deliver the dose to the intended treatment volume. Traditionally, patient set-up is carried out by aligning markers on the patient's skin (which are drawn when the patient is in the CT scanner) with orthogonal lasers in the linear accelerator room [7, 8]. This provides basic external set-up guidance, and internal set-up errors are traditionally measured and corrected using megavoltage portal film measurements or an electronic portal imaging device (EPID) [9].

The overall goal of radiotherapy is to maximise dose to cancer cells, while sparing the healthy tissue surrounding it to the greatest extent possible. To aid in the realisation of this goal, most patients are prescribed a regime of fractionated treatment. Instead of delivering a single, large dose in an attempt to achieve tumour control, the patient is subjected to a number of smaller dose fractions, typically five per week over six or seven weeks, depending on the nature of the tumour [6]. The rationale for this lies in the difference in recovery properties of healthy tissue cells and cancer cells [10]. As a result, patient set-up needs to be carried out not once, but for every treatment fraction.

1.3 Image-Guided Radiotherapy

Over the last decade, there has been a strong focus on the development of highly conformal radiotherapy techniques which are able to accurately deliver regions of high dose to tight volumes within the patient [5, 11]. Three-dimensional conformal radiotherapy (3DCRT) and intensity-modulated radiotherapy (IMRT) [12] are examples of such technologies. These techniques require advanced strategies for patient set-up on the linac to ensure the region of high dose is delivered to the correct volume.

This need led to the development of image-guided radiotherapy (IGRT) [5]. The aim of IGRT is to acquire anatomical information to use for further position guidance after room laser set-up has been carried out. This information is compared to the planning CT, so that alignment errors can be detected and corrected, if necessary [13]. Cone-beam computed

tomography (CBCT) is currently the most commonplace method of IGRT. CBCT-equipped linear accelerators have a CBCT unit mounted orthogonally to the linac gantry, and can reconstruct a volumetric CT image from a single rotation of the gantry around the patient [14].

The CBCT scan shows the position of soft tissue and bony anatomy inside the patient, and is compared to the planning CT. The degree of error in the initial patient set-up becomes apparent, and the treatment couch is translated and rotated [15] to minimise the geometric offset of the patient from the desired position [14]. This is known as a rigid body transformation, as the entire patient is shifted without applying deformations to individual parts of the body.

With the amount of information that has become available with in-room volumetric imaging for treatment, it has become apparent that patient position registration cannot always be achieved using only rigid-body transformations [16]. Over the course of fractionated treatment, it is common for changes to patient anatomy to occur due to, for example, weight change, organ movement and response of the target volume to treatment. More commonly though, rigid-body transformations are challenged by variation in patient pose. Room laser set-up only allows the patient to be aligned in the transverse plane, and deformations outside of this plane can impact the quality of treatment by contorting anatomy closer to the tumour.

An example of this effect can be seen in prostate treatments. Positioning the legs so that they are flat or raised on the treatment couch can make enough of an impact on prostate registration to affect the treatment plan [17]. These sort of deformations cannot be corrected with rigid-body alterations alone [18, 19]. Breast cancer treatments are also likely to suffer from non-rigid deformations [20, 21] (due to, for example, incorrectly positioned arms), as are head and neck tumours [22].

In this project, a system was devised that allows the radiation therapists to visualise patient pose variations during the set-up procedure, so that they can be corrected prior to rigid-body corrections with CBCT. The system was developed using a technology called "Augmented Reality."

1.4 Augmented Reality

Augmented Reality (AR) is a computer visualisation concept that involves the addition of virtual objects into a real-life scenario. It is similar to the better-known concept of virtual reality, in which the user's vision is completely immersed in a virtual world. AR, however, does not replace the real world with a fabricated one. Instead, it supplements the real world by taking a live image and superimposing onto it virtual objects that appear to coexist with the real environment [23, 24]. These virtual objects are usually added with the intention of providing additional perceptual information that would not be apparent without augmentation.

The commonly accepted definition of an AR system was outlined in a paper by Azuma in 1997 [23]:

- It combines both computer-generated and real objects.
- The user can interact with the system in real-time.
- The position(s) of the virtual object(s) in real-space are registered in 3D.

Augmented reality systems have enormous potential in the medical industry, due to their ability to provide intuitive support for both diagnosis and treatment by fusing 3D scan data with a view of the patient [25]. Already, AR has been successfully implemented in a number of medical applications [26–28], including some in the field of radiation oncology [29–32]. There has, however, been little focus on the development of AR technology to assist patient position guidance.

With the ever-increasing precision in dose delivery offered by modern radiotherapy methods, there has recently been a strong focus on the development of position guidance techniques to complement them. Several new technologies have surfaced in recent years for position guidance, such as laser scanners to detect patient geometry [33], stereo-vision for patient surface imaging [34], and detection of infrared optical markers on the patient's surface [35]. Although AR is a relatively young technology, the interest in positioning systems within the field of radiation oncology would make an AR guidance system valuable.

1.5 Objectives

Using augmented reality, the aim of this project was to develop a system to aid in position guidance during patient set-up in radiotherapy. An overview of the operation of this system is described in Figure 1.2. A video camera inside the treatment room obtains a live image of the linear accelerator, which is displayed on a monitor (Figures 1.2a and 1.2b). On this image, a 3D model of the patient's external surface is superimposed onto the linac couch, correctly positioned for treatment (Figure 1.2c). The radiation therapists observe the image and align the patient with the virtual contour on the screen (Figure 1.2d). When this is achieved, the patient's pose is correct.

As has been mentioned, there are a number of well-established position guidance systems that are already widely in use in radiotherapy departments. This AR system was not intended to replace any of these. Rather, it was intended as a supplement to these systems, to provide additional guidance without being invasive to the procedure as a whole.

The system was developed so that it would be able to offer the following advantages:

- When the patient has been aligned using room lasers, misalignments outside of the transverse plane of these lasers are usually not apparent. This system allows the surface of the entire region of interest (the region of the patient that is encompassed by the planning CT scan) to be aligned visually.

- Corrections that are applied after cone-beam CT set-up verification are rigid-body only (the couch itself is shifted while the patients retain their pose). With this system, non-rigid deformations become visible to the radiation therapist and can be corrected before taking a CBCT scan.

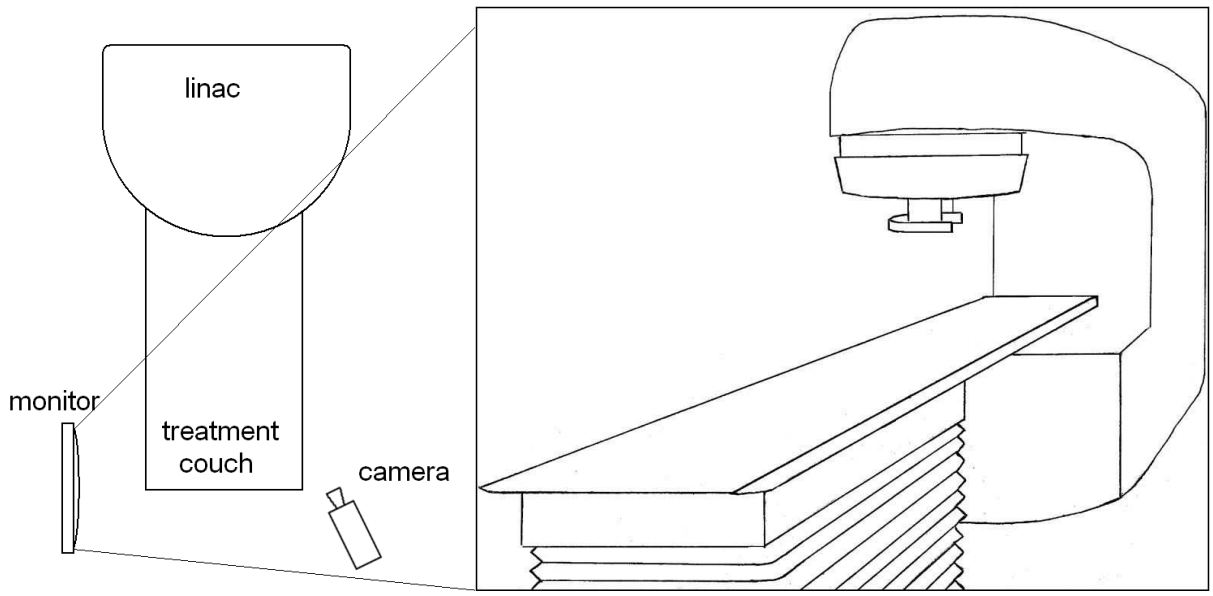
- Even after set-up is complete and the patient is in position, there is still the possibility of patient movement throughout treatment. This system does not only assist in the set-up process, but can also remain in operation throughout treatment. It then allows the radiation therapist to see if the patient deviates from the set-up position.

- Over the course of fractionated treatment, changes to patient geometry are likely to occur as their weight changes. The magnitude of these changes can be visualised directly with this system.
- It can be set up without any modification to the treatment equipment.

Throughout this project, there was a strong focus on optimising the precision and visualisation methods of the system, and performing tests to measure the achievable set-up accuracy. However, the main objective was to investigate the overall feasibility of augmented reality as a tool for position guidance, and to identify the issues to address before the system could become clinically viable.

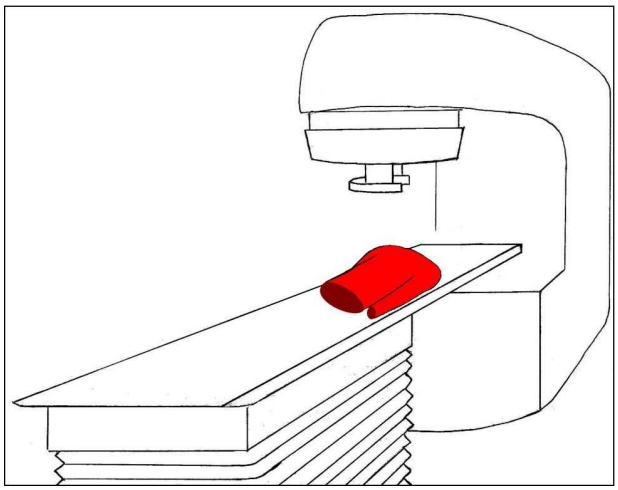
1.6 Contributors

This project has been a joint venture between the University of Canterbury Department of Physics and Astronomy and the Human Interface Technology Laboratory New Zealand (HITLabNZ), who are specialists in the field of augmented reality. HITLabNZ provided the equipment and the tracking and calibration software necessary to develop the AR position guidance system, and offered their expertise in the field of AR. My own contribution within the Department of Physics and Astronomy, was to apply the knowledge and tools provided by HITLabNZ into the field of radiation oncology to devise the position guidance system.

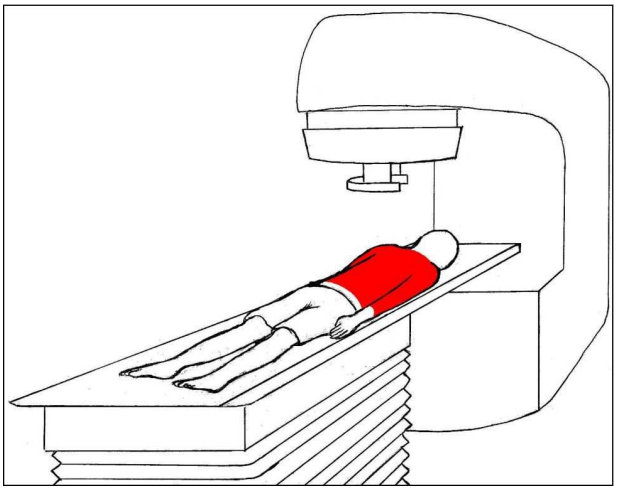


(a) External view of the system.

(b) The monitor display.



(c) A virtual image of the patient's external contour is superimposed over the camera image, correctly positioned on the treatment couch.



(d) The patient is placed on the couch and aligned with the virtual contour.

Figure 1.2: Visual outline of the AR procedure.

Chapter 2

System Development

In order to create the set-up guidance application, pre-existing augmented reality software libraries were used. The challenge lay in applying this software in an appropriate way, to make it suitable for use in patient set-up. Doing this required the acquisition of patient data for display, and the use of various apparatus and computer programs to perform co-ordinate registration and calibration. This chapter describes this development process and the challenges that were encountered, as well as the methods that were applied to address them.

The following equipment was used throughout development, and during system testing later on:

- Two Panasonic 2.3 megapixel digital video cameras (model no. NV-GS180EB).
- Two camera tripods (maximum height 175 cm).
- A desktop PC with a 17 inch, 1280 by 1024 pixel monitor, NVIDIA GeForce 8600 GT graphics card and Intel Core 2 Duo CPU E6750 @ 2.66 GHz.
- A Dell laptop, with a 17 inch, 1600×1200 pixel display, NVIDIA GeForce Go 7950 GTX 512 MB graphics card and Intel Core 2 CPU T7200 @ 2.00 GHz (this was used when the system had to be taken outside of the laboratory).

2.1 3D Patient Data

In order to create the AR program, a 3D model of the patient's external contour had to be obtained. This model had to reproduce the patient's pose from the planning CT scan, so the obvious source from which to obtain the model was the CT scan itself.

A CT scan consists of a series of transverse slice images of the patient's anatomy (Figure 2.1). The distance between each slice varies depending on the purpose of the scan, but it is usually between 2.5 and 10 mm. Combined in a series, the slices provide volumetric data of the patient's anatomy. When the scan is complete, the files are typically transferred and stored in the DICOM (Digital Imaging and Communications in Medicine) data format [36], which is a universally accepted standard for the storing and transfer of images between medical equipment. DICOM objects contain a range of data attributes, including patient information, scan type, scan parameters, image parameters and the pixel data that specifies the image itself.

The treatment planning system (TPS) is used to draw contours around the edges of the volumes of interest (VOIs) in each of the CT slices (Figure 2.1). These usually include the patient's external surface, any organs at risk, the clinical target volume (CTV– the gross

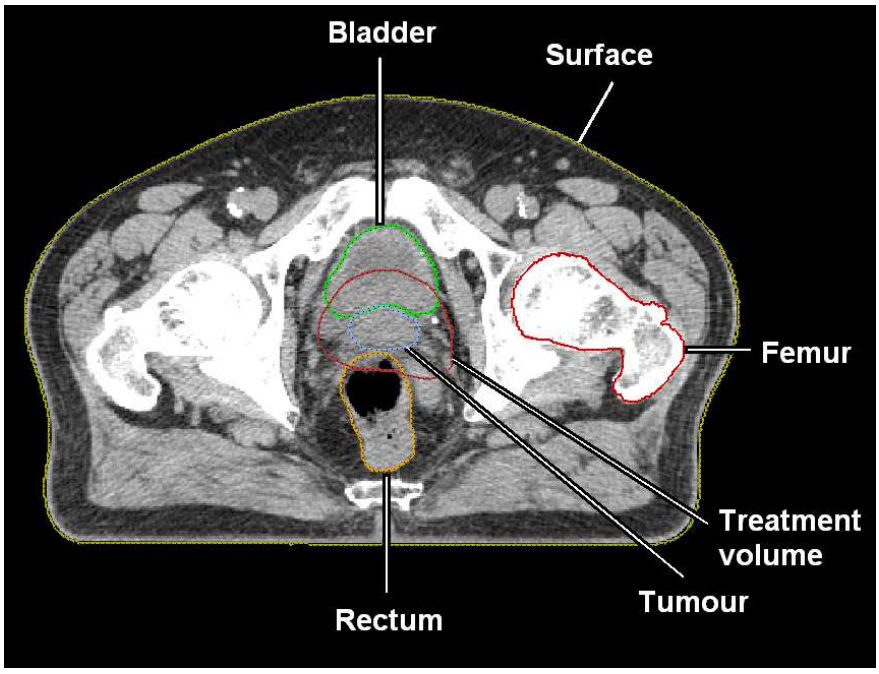


Figure 2.1: Pelvic CT scan of a patient with prostate cancer.

volume of the tumour) and the planning target volume (PTV– the total volume to be treated, including margins to account for patient motion and set-up errors). The contour lines are comprised of a number of vertices around the edges of the structures, which are connected by small line elements.

The contour information is stored as a separate "radiotherapy structure" (RS) DICOM object. This does not contain any actual image data, however the metadata associated with it contains the set of vertex coordinates (the pointset) for each of the structure contours. These pointsets are retained so that the TPS can generate surface renderings to provide a 3D reconstruction of the patient's anatomy. Note that treatment planning systems generally do not include the option to export surface reconstructions of the pointsets to separate 3D model files.

It was therefore necessary to write a program which could read the contour vertices from the DICOM-RS file and generate a 3D model in an appropriate file format. This program was written in MATLAB [37], using the Image Processing Toolbox [38], which has a number of functions that are useful for reading DICOM objects and their associated metadata.

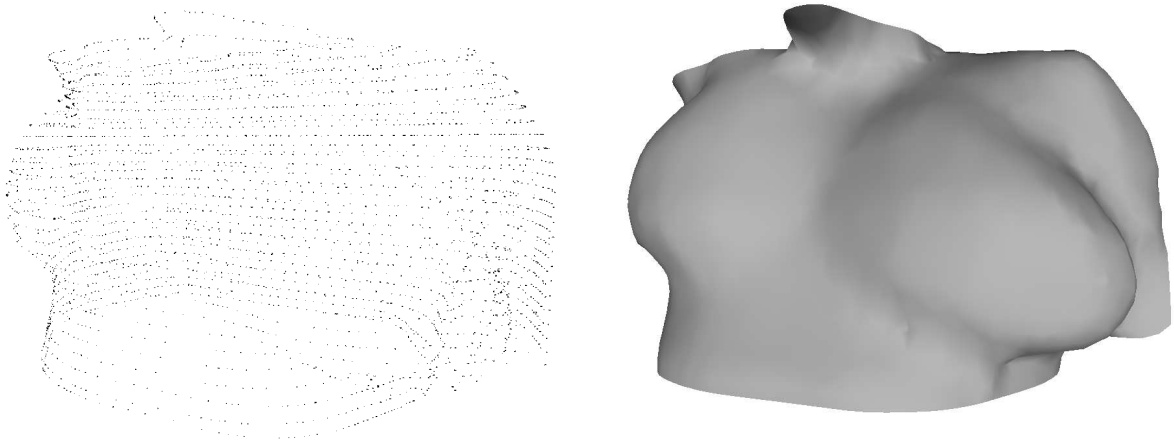
The MATLAB program prompted the user to select a DICOM directory, and it loaded the DICOM-RS file. The user selected the appropriate contour (i.e. the patient's external surface), and the program proceeded to read the vertex coordinates associated with it. It then wrote the vertices into two separate 3D model files: the first was a pointset and the second a wireframe of the contour. These two models were stored as Wavefront Objects and AutoCAD Drawing Exchange Format (DXF) files respectively (in depth descriptions of these file formats can be found in [39] and [40]).

DICOM-RS files only contain the vertices of the patient contours. They do not contain any information regarding how the vertices should be connected or how faces should be arranged. This meant that the Wavefront Objects that were generated only contained vertices, and lacked connecting lines and faces. These objects (pointsets) were insufficient for visualisation (Figure 2.2a). The surfaces of the models therefore had to be generated with a surface reconstruction algorithm (Figure 2.2b).

An open-source application called MeshLab [41] was used for this purpose. Meshlab

uses a "ball-pivoting" algorithm [42] for surface reconstruction of pointsets. It also contains algorithms which can be used to smooth the surface of a 3D model, although smoothing was not used excessively in this project, so that the original contour derived from the planning CT was preserved.

The wireframes were generated by reading the vertices from the DICOM-RS object, writing them to the DXF and connecting each subsequent vertex with a straight line (Figure 2.3). While the surface-rendered Wavefront Objects were used as the primary model for visual augmentation, the wireframes were used to improve the visualisation by providing the user with enhanced depth perception (see section 3.2.2).



(a) The vertices from the patient surface contour are exported into a Wavefront Object pointset.

(b) The patient contour after surface reconstruction has been performed.

Figure 2.2: Wavefront Object obtained from a breast treatment scan.

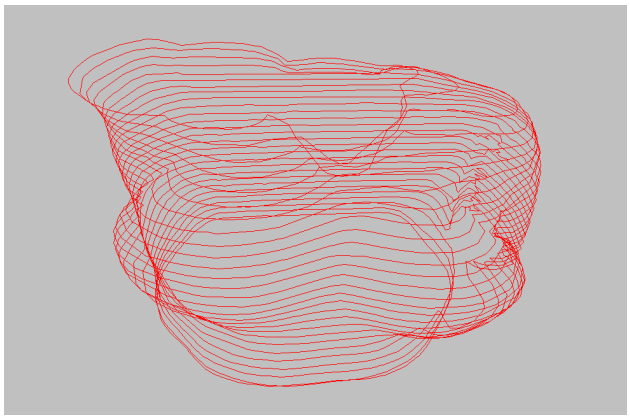


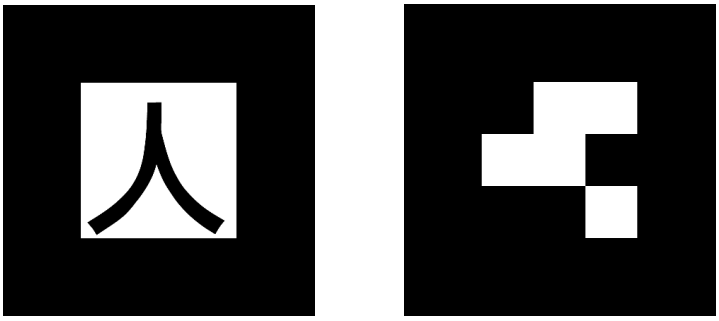
Figure 2.3: Wireframe 3D model from a breast treatment scan.

2.2 Software

The next step was to create an augmented reality application which obtained a live image from a video camera and superimposed the patient surface model onto it, so that it gave the impression of existing in the real environment. HITLabNZ provided the software that was necessary for this project, namely OSGART [43], a C++ library supporting the development and prototyping of augmented reality applications. OSGART combines the high-performance 3D graphics library "OpenSceneGraph" [44], with the well known augmented reality tracking library "ARToolKit" [45][1].

ARToolKit uses computer vision tracking algorithms to deduce the location and orientation of the video camera in real-time. To do this, in every frame of the video input, it searches the image for certain fiducial tracking markers, which are easily identifiable black squares. The software can distinguish one marker from another, as each of these squares contains a unique pattern (Figure 2.4).

ARToolKit includes a range of patterns that can be printed (Figure 2.4a), or the user can generate their own custom patterns. It also supports "binary matrix" markers (Figure 2.4b), which provide more reliable pose estimation of the 3D objects that are added to the scene. These markers are intended for applications where aesthetics are unimportant, such as medical applications, and so were used in this project.



(a) Tracking marker pattern included with ARToolKit.

(b) 3×3 binary matrix tracking marker.

Figure 2.4: Examples of fiducial tracking markers used with the ARToolKit tracking library.

[1]The version that was used in this project was ARToolkit Professional 4.3, which includes more accurate pose-estimation algorithms than the open-source version.

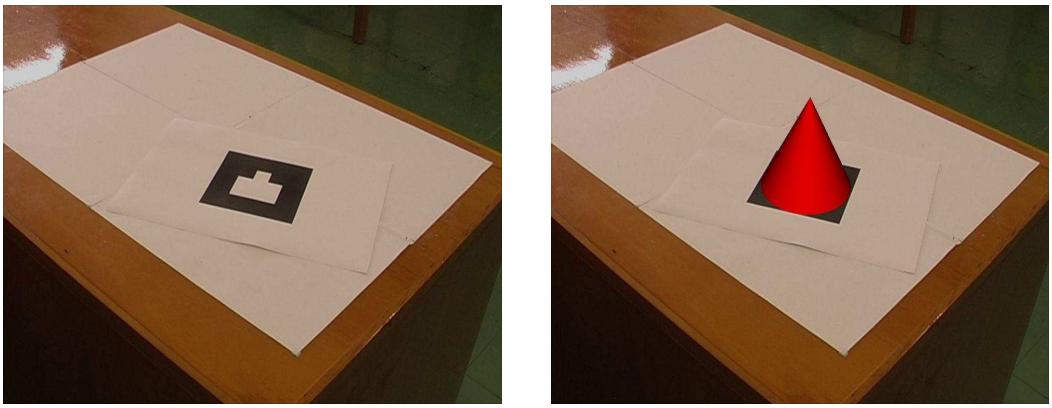
The patterns in the binary matrix markers were 3×3 grids of black and white squares, surrounded by a black border. The geometric constraints in the marker patterns made them easier to detect, and their use resulted in a more accurate and stable system.

Upon detecting the markers (of known size) in the video image, the ARToolKit tracking algorithms could calculate the geometric transformation matrices that relate the marker coordinates to the camera coordinates:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} V_{11} & V_{12} & V_{13} & W_1 \\ V_{21} & V_{22} & V_{23} & W_2 \\ V_{31} & V_{32} & V_{33} & W_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} = \begin{bmatrix} & \mathbf{V} & & \mathbf{W} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix}$$

where X_c, Y_c and Z_c describe the coordinates of the camera, and X_m, Y_m and Z_m describe the coordinates of the tracking marker. $\mathbf{T}$ is the transformation matrix that relates the two, and is comprised of the matrix $\mathbf{V}$ (which describes the rotation between the two coordinates sets) and the vector $\mathbf{W}$ (the translation between them). A detailed description of how ARToolKit estimates the transformation matrix $\mathbf{T}$ can be found in [46]. With the position of the camera relative to the markers known, the software could superimpose 3D objects into the scene, in positions relative to the markers (Figure 2.5).



(a) In the video image, a square tracking marker is visible.

(b) Tracking software is enabled. A 3D cone is displayed on top of the marker.

Figure 2.5: ARToolKit tracking example

2.3 Display System

One of the first things that had to be considered for the AR system was the way in which visual information was displayed to the user. There is a vast range of approaches that can be taken when designing an AR display system [47], several of which have been investigated in medical environments [26–28]. In this project, as with all AR systems, it was important to design a display system with intuitive usability so that it did not convolute workflow in practice. However, the simplicity of the system should not come at the cost of precision and accuracy, thus undermining its effectiveness.

The following display systems were considered and evaluated for their practicality in this project:

Head-mounted Displays

Head-mounted displays (HMDs), currently exist in two main forms: video see-through [48] and optical see-through [49] configurations. Video see-through systems consist of a camera/display combination mounted directly in front of the eyes, while optical see-through displays usually have semi-transparent mirrors with virtual objects reflected off them from a projector or monitor in the headset. Video see-through displays are fairly reliable in terms of their registration abilities, since they operate on the same principle as any other camera/display AR system. However, they tend to reduce the perceptive abilities of natural eyesight; viewing the world in a display as seen by a camera is not the same as looking through one's own eyes. Optical see-through displays allow the user to retain their natural vision, but require calibration to account for parallax between the eye and the display.

HMDs eliminate the need for the user to glance from the real environment to a display in order to observe the augmented world. Thus, overall they have the advantage of immersing the user directly into the augmented world, providing direct 3D interaction. They also require very little in the way of additional equipment within the treatment room, since a camera and display are not needed in the environment.

Despite their advantages, it was decided that an HMD system was not practical for this

project. For the system to be clinically viable it had to be as discrete as possible. The aim was to create it as an optional tool for the radiation therapist, one that could be accessed with minimal disruption to workflow. Additional cumbersome equipment such as a head mounted display would have been an inconvenience to the user.

AR Windows

Essentially, an AR window is a semi-transparent window or mirror that is placed between the user and the environment, so that the augmented world is visible when the user looks through it. Virtual objects can be added to the scene by projecting them onto the window, or by reflecting a display from a nearby monitor.

With AR windows, the position of the user must be tracked so that transformations can be applied to the virtual objects to account for the viewing angle. This and the fact that an AR window needs to be placed directly in front of the viewer rendered them unsuitable for this project, since an instrument that acts as an obstruction between the radiation therapist and the patient would be likely to disrupt the set-up process.

Hand-held Displays

Hand-held monitors generally consist of a small flat-panel screen displaying a video image from a camera attached to their back, and a handle for user interaction. Augmentations are applied to the image on the monitor, so that they act as "virtual lenses". They can be moved freely in space, and so provide a convenient and intuitive interface, without being too invasive to the user's sensory perceptions.

In this project, a hand-held display would have been impractical because of the limited screen size, which would need to encompass the entire body of the patient. It would also occupy an otherwise free hand of the user.

Projection Displays

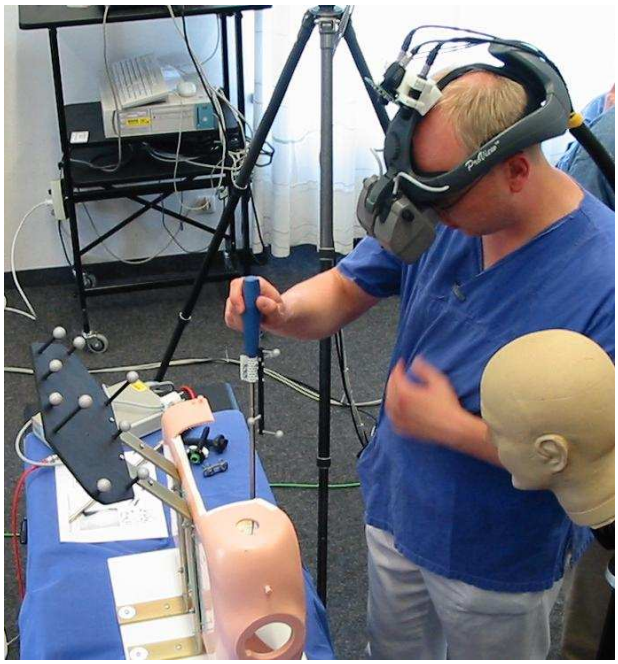
Projection displays use a projector to cast augmentations directly into the real environment. These have the distinct advantage of directly enhancing the user's vision in an intuitive

way, without necessitating any additional equipment for visualisation. However, since the augmentations are only visible on the surface of solid objects, they rely on the placement of objects within the environment, and essentially only provide 2D information. There is also the matter of occlusion generated by the presence of people or instruments between the projector and the patient, thus limiting the freedom of movement of the user. Such a system is better suited to an application that generates markers for targeting purposes, for example, on the surface of a patient or on medical equipment.

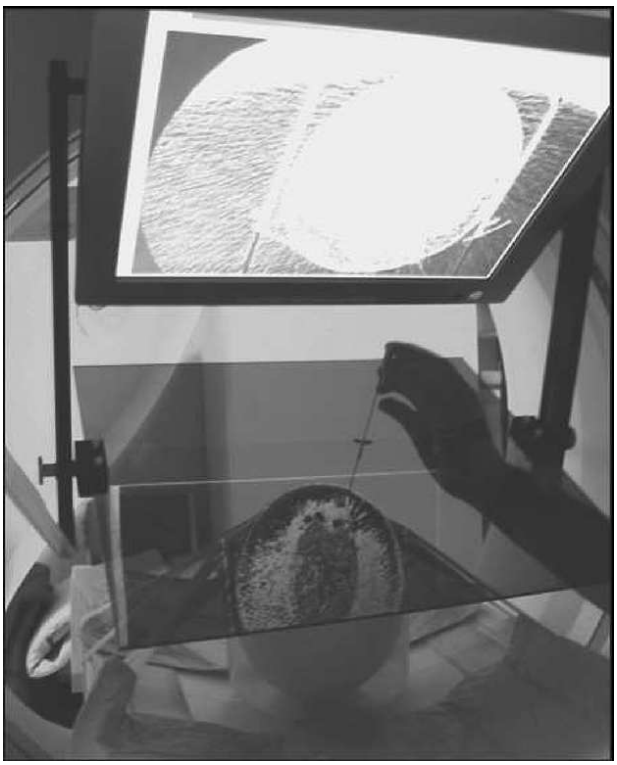
Augmented Monitors

An augmented monitor is a comparatively simple display system for an AR system, requiring a minimum of a camera to obtain the live image, processing equipment and a monitor for the display. The traditional augmented monitor set-up was deemed the most appropriate arrangement for this project. Such a system could operate in the periphery of the radiation therapist, so that the augmented display would be available on the monitor whenever they required visual guidance, yet it would not encroach on existing patient set-up methods. An augmented monitor also offered the advantage of being visible to any radiotherapy staff contributing to the procedure, whereas a display system such as an HMD would be limited to a single user.

The main disadvantage of augmented monitors is that they do not project augmentations directly into the user's vision like HMDs or projectors, so they do not provide the same degree of 3D interaction. The user has to glance back and forth between the scene in front of them and the monitor, which involves a supplementary cognitive effort to mentally match the spatial relationship between their own vision and the monitor display. Augmented monitors also necessitate additional equipment (i.e. the camera and monitor) within the treatment room. However, the acceptance of newly proposed technology is challenging in the medical field, as it requires usablility testing and the adaption of staff to unfamiliar procedures. It therefore made sense to prototype and evaluate the interest of AR patient set-up guidance with a common and well-known display system.

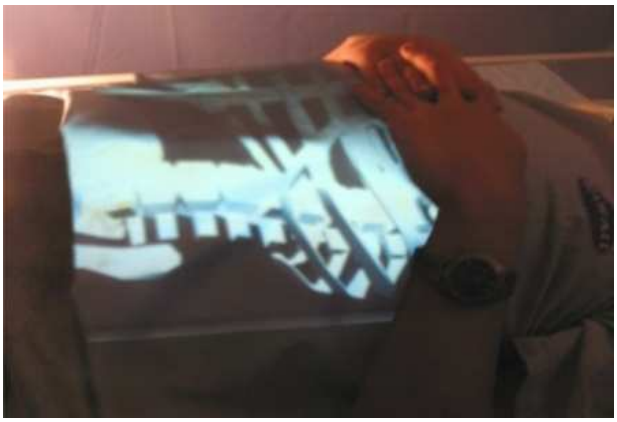


(a) HMD developed for computer assisted needle insertion [50].



(b) AR window for needle insertion guidance in a CT scanner [51].

(c) Hand-held AR device to visualise anatomy [52].



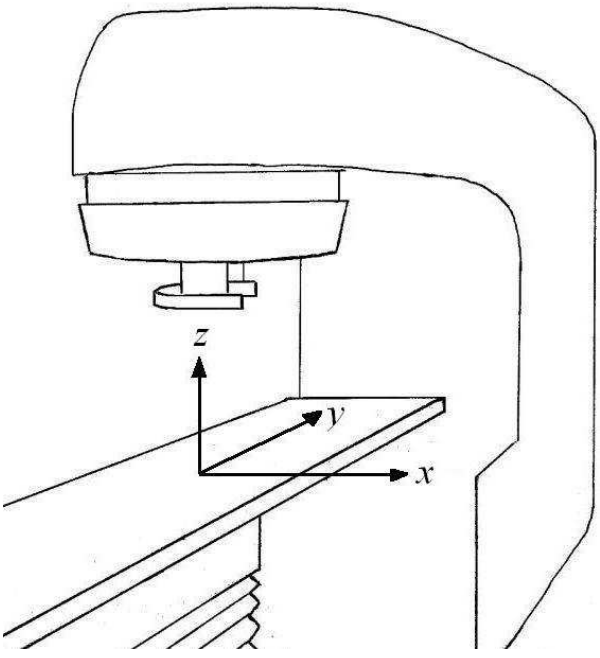
(d) Projector-based AR, allowing visualisation of skeletal structure [52].

Figure 2.6: Examples of AR in the medical industry.

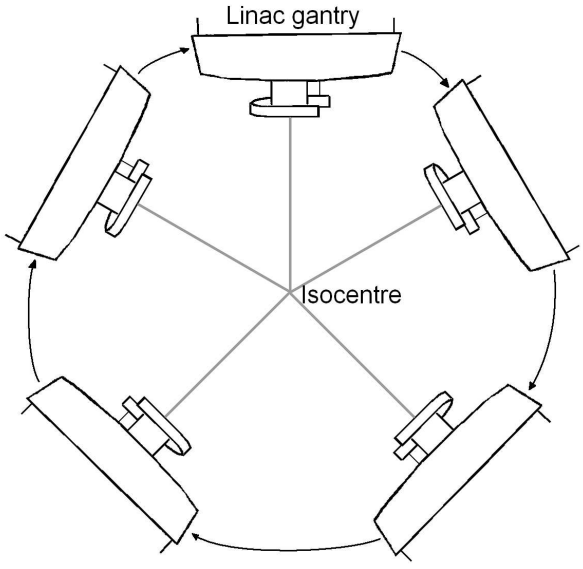
2.4 Registration

Registration is the process of bringing two or more images (2D or 3D) into spatial alignment in a way that fuses the useful data in each, so that they complement each other [53]. In this case, the 3D model of the patient's external contour had to be registered with the live image of the treatment room so that it was aligned in the appropriate position on the linac. A mathematical model of the treatment room and the spatial relationships between different relevant components (the camera, linac and patient model) had to be created.

To fully grasp the technicalities of this process, first one must understand the coordinate systems involved in a radiotherapy procedure. A set of cartesian coordinates associated with the linear accelerator can be defined (Figure 2.7a), the origin of which is located at a point in space called the "isocentre." In order to deliver x-ray beams to the patient from a variety of angles, the linac gantry is able to rotate 360° around the treatment couch in the plane shown in Figure 2.7b. The point where the radiation beams converge is the linac isocentre. This point remains fixed in space in every treatment; the gantry never shifts out of the xz plane, rather, the treatment couch itself is translated to orient the patient with the isocentre.



(a) A system of cartesian coordinates is defined for the linear accelerator.



(b) As the linac gantry rotates around 360°, the radiation beams converge at a point known as the *isocentre*, which acts as the origin.

Figure 2.7: The linear accelerator coordinate system.

During treatment planning, the point where the radiation beams should intersect inside the patient is chosen. This point is known as the *treatment* isocentre, and is marked in the planning CT dataset. In most cases, it is located within the volume of the tumour itself. The treatment isocentre should, ideally, align with the linac isocentre when the patient is set up for treatment.

The two coordinate sets that had to be registered into spatial alignment therefore already had a common reference point; the treatment isocentre within the patient had to align with the linac isocentre. This left the question of how to make the tracking software recognise where the linac isocentre was located in the camera image. The obvious answer was to attach fiducial tracking markers within the camera's vision. If each marker were positioned so that its location and orientation relative to the isocentre were known, the appropriate parameters describing these locations could have been entered into the system. The program would have then been able to deduce where the isocentre was located within the image, from the positions of the markers that were visible.

Two problems with this approach became immediately apparent. First, the isocentre is a virtual point in space, and there was no way of either attaching markers directly to it, or to accurately measure the distance to any marker that was positioned nearby. The distance to a nearby marker could have been approximated, but measuring from an invisible point in mid-air would have had a large uncertainty associated with it. This would have been unacceptable, since accuracy in patient positioning is of utmost importance.

The second problem was that for the patient contour augmentation to remain stationary in the camera image, the fiducial markers had to remain stationary as well. An ordinary linear accelerator bunker does not contain any suitable flat surfaces that remain still for the entire treatment process. The surfaces that were available are shown in Figure 2.8, and were as follows:

(a) The most obvious position was on the treatment couch. The surface of the couch is flat, and since the markers would be close to the patient, the augmentation in the video image would have been fairly stable. However, during the treatment process, the couch itself is translated and rotated to bring the patient into alignment with the

linac isocentre. The virtual object on the monitor would therefore shift in unison with the couch. It would have been possible to position the augmentation so that it lay on top of the couch, but this would mean that the user would essentially be aligning the patient relative to the couch, not the isocentre.

(b) The gantry of the linear accelerator was unsuitable for the same reason. During treatment it rotates around the patient to deliver the radiation beams from all the necessary angles, so the virtual object would rotate with it.

(c) The walls and floor in the linac bunker remain stationary throughout treatment, but their distance from the patient raised some problems. The further a virtual object is to be displayed from a reference point (i.e. a fiducial marker), the more extrapolation is required in order for the system to calculate the geometric transformations. This method would have caused a great deal of uncertainty in the position of the virtual object on the monitor. Even if the sacrifice in precision had been deemed acceptable, the problem of measuring the exact relative coordinates of the isocentre to each of the markers would have been considerable.

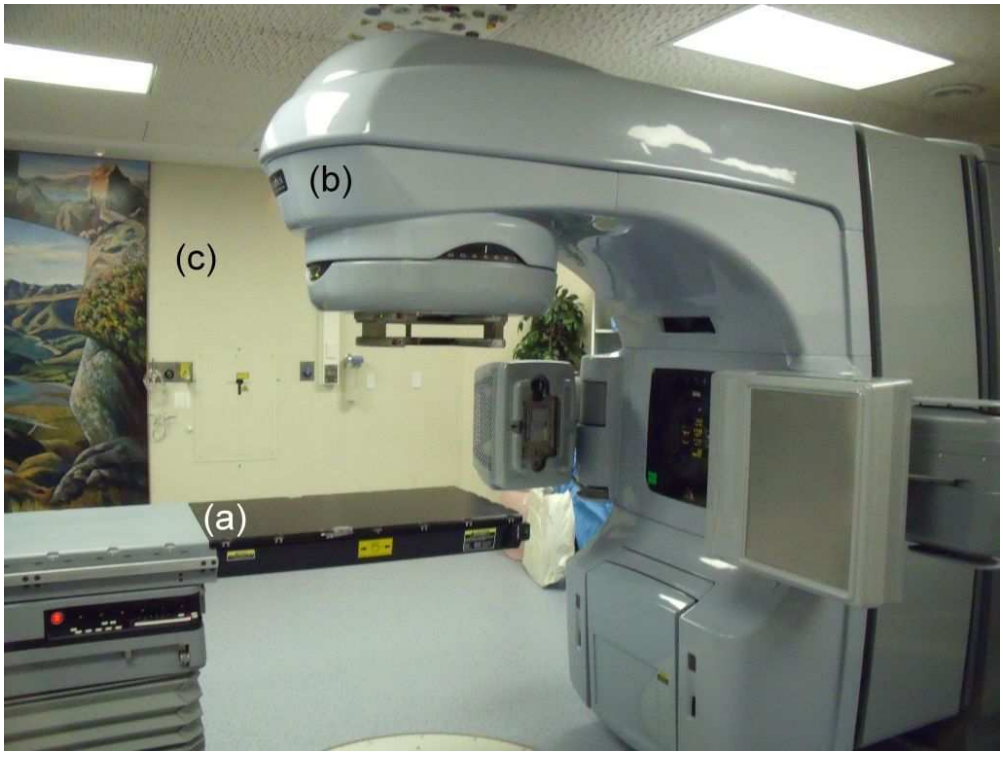


Figure 2.8: Locations in the linac bunker that were considered for fiducial tracking markers.

In answer to these problems, a novel technique was developed to precisely register the coordinate systems. The room lasers that are used for initial patient set-up intersect at the linac isocentre, so if an object with fiducial tracking markers attached to its surface were placed on the treatment couch and aligned with these lasers, registration could be easily performed.

A small cube of side length 92 mm was constructed out of solid perspex for this purpose (figure 2.9). Vertical and horizontal grooves, each with a width of 1 mm, were cut through the centre of each face so that, when the treatment couch was adjusted to the appropriate height, it could be aligned with the fan lasers in the linac room. When aligned, the cube centre corresponded to the linac isocentre.

Twenty fiducial tracking markers were glued to the so-called "registration cube," with four on each face, apart from the underside. The positions of each of these markers relative to the centre of the cube were entered into the system, so that when viewed by the camera from any angle, the system could detect the markers and thus register the origin of the coordinates to the centre of the cube (i.e. to the isocentre).

When the cube was aligned with the lasers, the coordinate system that had been established could be locked into place by pressing the "enter" button, disabling tracking of the fiducial markers. The cube could then be removed from the couch, so that the patient could be positioned and treatment could proceed.

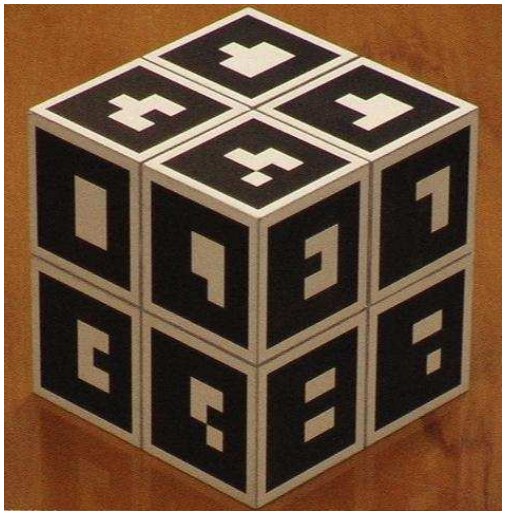


Figure 2.9: The registration cube.

As an alternative to room-laser alignment, it was possible to align the registration cube with the linac isocentre using cone-beam CT. This required a CT scan of the cube prior to implementation of the system, which was obtained at Christchurch Hospital with a GE 64 slice CT scanner. The cube could then be aligned as above and a CBCT scan obtained. Comparison of the CT data with the CBCT data would reveal any offset of the cube from the linac isocentre, and the couch could be translated and rotated to compensate.

To aid the CBCT alignment process, small holes (4 mm in diameter) were drilled into the centres of four of the cube faces, in which steel screws were affixed, flush with the cube surface. These screws were to act as landmarks in the CT and CBCT volumes, using the sharp contrast between them and the surrounding perspex.

2.5 Tracking

In augmented reality, tracking refers to the real-time pose-estimation of a virtual object [26]. In other words, as the position of the camera is changed, or as certain objects move in real-space, geometric transformations are applied to the virtual objects in real-time so that their locations and orientations are updated to match the dynamic environment. As described in section 2.2, ARToolKit performs tracking by detecting fiducial markers within the space that is visible to the camera.

With the registration cube and the tracking markers removed from the camera's field of view, tracking could no longer occur. This raised a problem with the registration cube method: after the registration process, the system was no longer able to apply transformations to update the position of the virtual patient contour on-screen. If any camera movement occurred, the augmentation remained fixed in place on the monitor, like an object that was attached directly to the camera lens.

This was not so much of a setback as it initially seemed. The use of augmented reality in this project differed to the norm in the medical industry. In many medical AR applications, the virtual objects need to be registered to the patient so that they are displayed in relation to the patient's current position. To achieve this, markers can be affixed directly to the patient,

and tracked using AR software [26]. In this project, the opposite was to be achieved; the objective was not to align the scan data with the patient, but to align the patient with the scan data. The virtual augmentation remained stationary as the patient's position was altered, thus making it apparent when the patient was misaligned.

The virtual patient contour was displayed with respect to the linac isocentre, which remains stationary at all times. As long as the camera remained stationary after registration with the cube, there was no need to update the pose of the virtual object. Removing the tracking ability of the system was therefore not a hindrance to its performance, so long as a suitable camera angle was chosen prior to registration. It was important though, to take care not to knock the camera during treatment, or to adjust any of the camera parameters (such as the zoom level).

2.6 Calibration

During the use of the system, some of the transformations between coordinate systems remained static, and could be determined through a calibration process. This included the initial positions of all the fiducial markers with respect to a common reference point, and correction factors to account for any distortions in the camera image.

2.6.1 Registration Cube Calibration

The twenty fiducial markers were placed with respect to the centre of the cube which, when correctly aligned, corresponded to the linac isocentre. It was necessary to enter into the system the parameters that determined how this point related to the arrangement of markers. In other words, the tracking software could recognise the markers, but without these parameters, it would not know where to display the virtual object in relation to them.

Every marker was assigned three parameters which described its position and orientation with respect to the centre of the registration cube: a 3×1 vector which stated its position, a 3×3 matrix which described its orientation, and a scalar which described the size of the marker.

Figure 2.10a shows the registration cube with only one of the tracking markers visible. The centre-point of the square marker, with respect to the centre of the cube, is described by the vector:

$$\mathbf{v} = \begin{bmatrix} 46 \\ -23 \\ 23 \end{bmatrix}$$

In other words, its centre is 46 mm in the x direction of the cube, -23 mm in the y direction and 23 mm in the z direction (Figure 2.10b).

Since markers were placed on five of the six cube faces, there were five possible orientations for the markers to be in. Without any information regarding these orientations, the system would have assumed that they were all facing upwards (in the z direction). They were therefore each assigned a transformation matrix which described their rotation with respect to the default, upwards orientation.

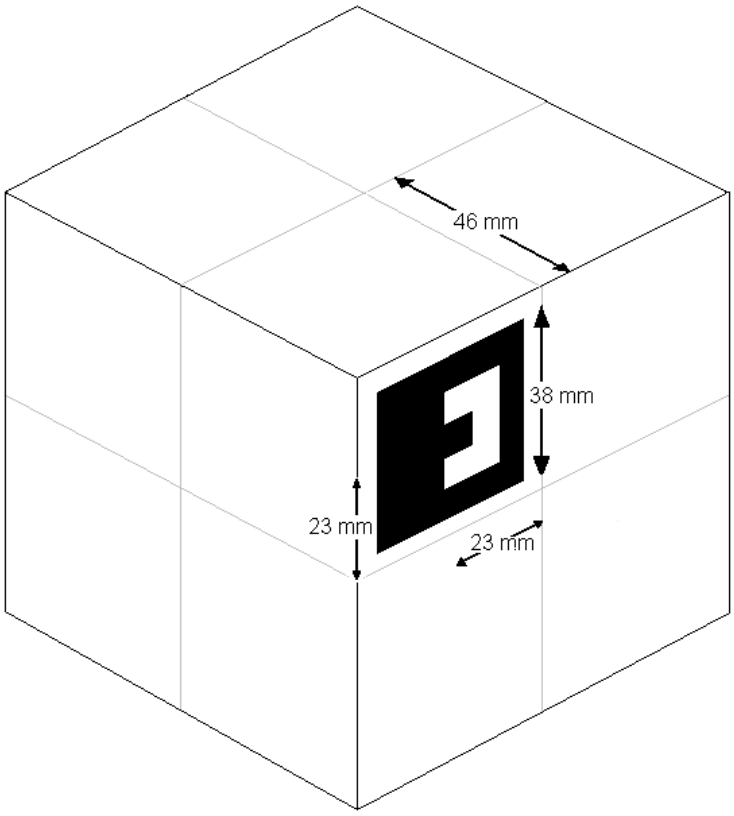
Euler's rotation theorem states that any rotation can be described by a composition of rotations around three axes, and therefore can be represented by a $3\times$ 3 matrix $\mathbf{R}$ operating on a vector, as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

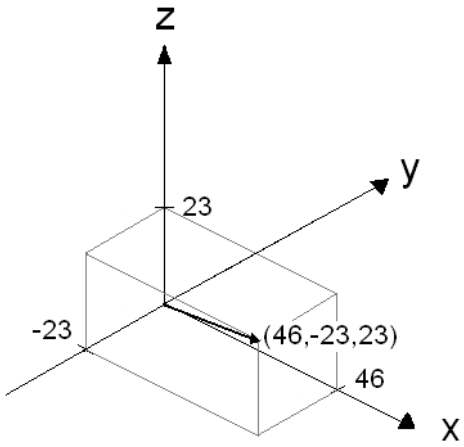
Fortunately, on the registration cube, the markers were all arranged orthogonally, i.e. the faces were at right angles to each other. In this case, the rotation matrices were therefore fairly easy to estimate analytically. As an example, the rotation matrix $\mathbf{R}$ for the marker in figure 2.10c was calculated as follows:

$$\mathbf{R} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} y + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} z$$

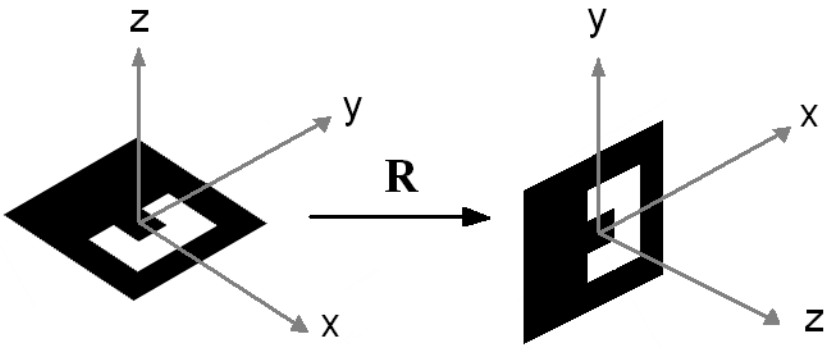
$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$



(a) Registration cube with a single marker positioned on one of the faces.



(b) The vector that describes the position of the centre of the tracking marker. The origin of the axes corresponds to the centre of the cube.



(c) The transformation **R** describes the rotation of the tracking marker from its default orientation to its orientation on the cube.

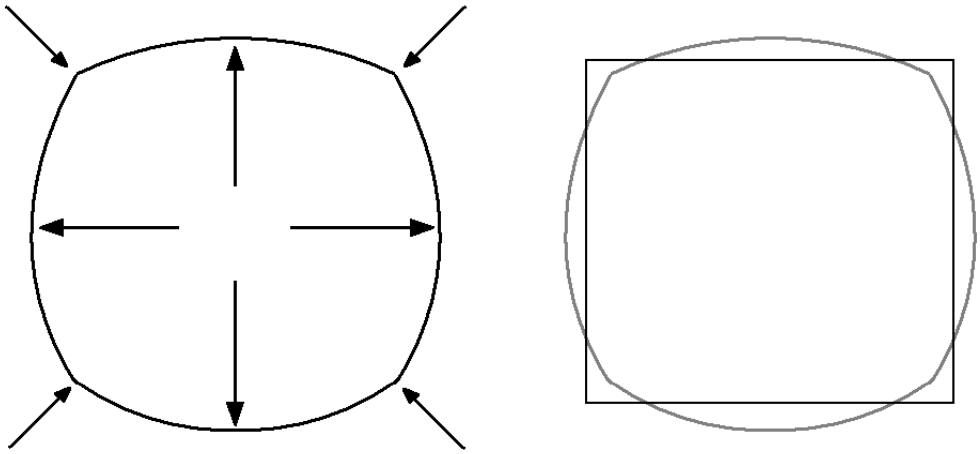
Figure 2.10: The registration cube.

Similar calculations could be performed for markers on each of the cube faces. The rotation matrices for the markers on the top face were simply the identity.

The final parameter in the calibration of the tracking markers was a scalar which determined its size. Without this parameter, the software would not have been able to determine whether a marker that appeared large in the camera image was a small marker close to the lens, or a large one that was distant. In the case of the square marker shown in figure 2.10a, this parameter was 38, meaning that the length of each of its sides (of the black part of the marker) was 38 mm. This was the same for all twenty of the markers, since they were all the same size.

2.6.2 Camera Calibration

Calibration of the video camera is a significant challenge in the creation of augmented reality programs. The image that is obtained by the camera is not necessarily the same as that which can be seen in real life. The camera lens is likely to have a distortion effect on the image, such as barrel distortion (Figure 2.11a), or some other effect that is intrinsic to the lens. This effect may be immediately obvious when viewing through the lens, or it may be so subtle that it is imperceptible to the unaided eye. Regardless, it is unlikely that a video feed obtained from a camera will be completely devoid of warping.



(a) Barrel distortion causes the edges of a camera image to be stretched outward.

(b) A virtual object that is placed in the scene (a square) does not suffer from the same distortion effects as the background image.

Figure 2.11: Distortion effect on a camera image.

Camera distortion affected the operation of the AR program in two ways:

- When the software searched an image for fiducial tracking markers, its perception of the markers was altered by camera distortion. For example, distortion tends to be more prominent towards the edge of the camera image. A marker that appeared in this region of the image would become skewed, and would come into conflict with markers that experience less distortion near the centre of the image. This would cause an error in the determination of the central reference point (i.e. the reference point would be offset from the centre of the cube).

- When the program placed the model in the scene, it was not distorted in the same way as the background. There was a discrepancy between the view of the background environment, and the view of the virtual object (Figure 2.11b). This would have been problematic while trying to align the patient with the 3D model of their external surface. If the real patient appeared warped by camera distortion, while the overlaid 3D contour retained its dimensions, the patient would have been positioned incorrectly.

To counter the effects of camera distortion, the ARToolKit library includes an application to calibrate the video camera, to allow AR programs to compensate for the intrinsic properties of the camera. In the open-source version of ARToolkit [54], the application has two steps, the first of which involves using the camera to take a number of snapshots of a planar calibration pattern (also provided with the library, for the user to print out). The user clicks on each of the twenty-four dots on the pattern (Figure 2.12a), and repeats for every image that is obtained (five to ten images are recommended). The program then deduces, from the positions of the dots, the parameters that describe the distortion of the camera.

The second step of the calibration determines the focal length of the camera. Another calibration pattern is printed, and again the user must take a number of snapshots with the camera. The program overlays horizontal and vertical lines on the screen and, using the keyboard, the user aligns these with the lines on the calibration pattern (Figure 2.12b) (the virtual lines are distorted using the parameters obtained in the first step of calibration). Again, this process is repeated for a number of snapshots. Before terminating, the application

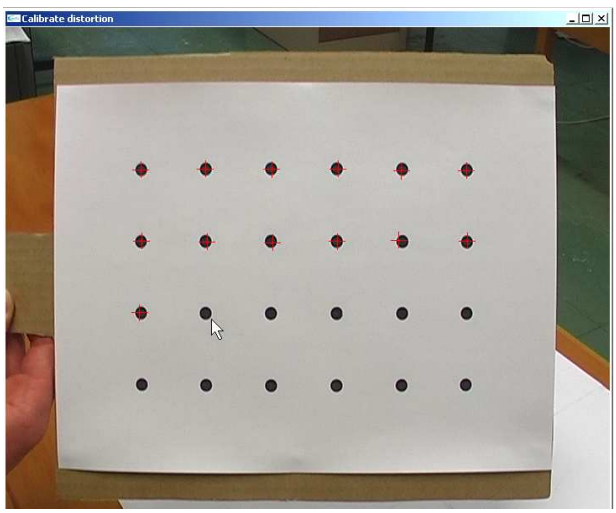
outputs a data file containing all the calibration information, which can be used with an ARToolKit-based application.

It was noticed that this calibration procedure suffered from two major shortcomings. The first was that it was extremely time consuming; the whole process took roughly fifteen minutes to complete, depending on how many images the user chose to obtain (a greater number of images produced better accuracy). If the system were to be used in a clinical setting, this would be a major disadvantage. Even if calibration were performed infrequently, say, as part of a daily or weekly equipment quality assurance routine, it would still incur a great deal of time pressure for hospital staff. The situation would be worsened if unexpected circumstances gave cause to recalibrate the system.

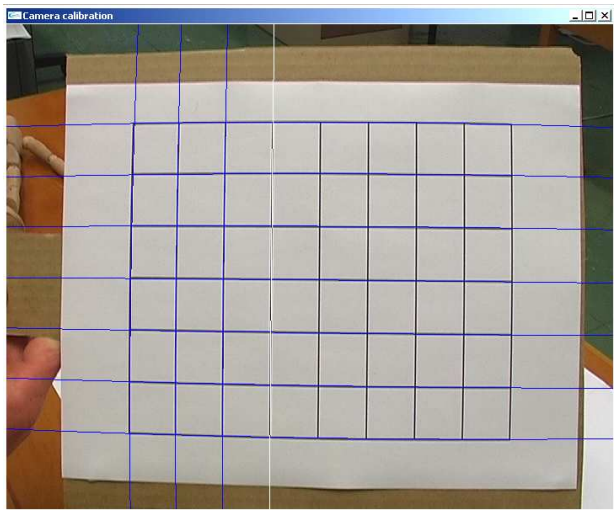
The second disadvantage of the calibration program was that it relied heavily on human input, which is prone to error. This would propagate to the treatment set-up itself, and may give cause to repeat the entire calibration process.

The two-step calibration program is included in the professional edition of ARToolKit (which was used in this project). However, a second, single-step application is also included. This program implements OpenCV [55], an open-source computer vision library, and is greatly simplified. As in the original program, the user prints a planar calibration pattern which, in this case, is a black and white chequered image of known size (Figure 2.13). Again, the user takes a number of snapshots of the pattern, from a range of angles and distances. The program is then able to automatically locate the corners of the squares in the images. The camera parameters are generated from this data.

This process was far less time-consuming, taking only one or two minutes to complete. It also minimised human input and the potential for error, so was better suited to this project.



(a) Step 1. The user clicks on each of the twenty-four dots in the image, and repeats for several images.



(b) Step 2. The user aligns virtual lines with grid lines in the image.

Figure 2.12: Two-step calibration process

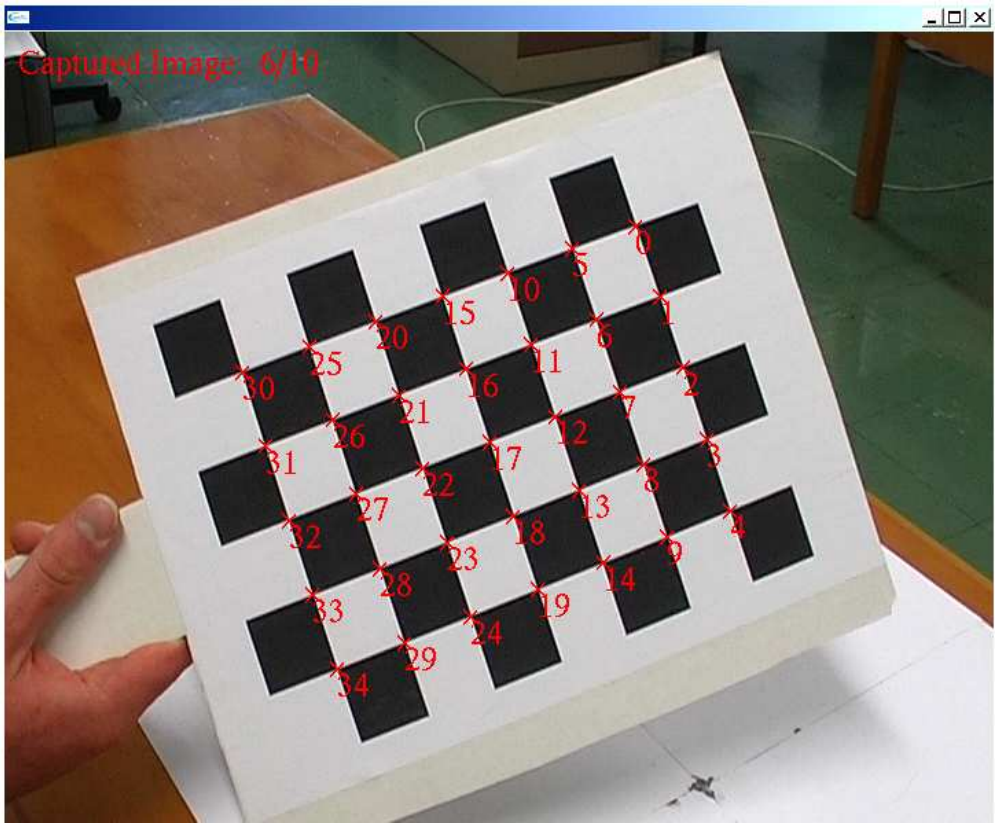


Figure 2.13: Single-step camera calibration program.

2.7 Coordinate System Relationships

The calibration parameters that were obtained were used to integrate all the relevant coordinate systems into a single "world coordinate system" (WCS), the unified set of coordinates in which the system operated. The world coordinates originated at the linac isocentre and were warped according to the camera calibration parameters to account for intrinsic camera distortion. Figure 2.14 shows each of the local coordinate systems, and illustrates how they were unified into the framework of the world coordinate system. Each arrow represents a transformation that combines one set of coordinates with another. The relationships are explained in Table 2.1. (Tuceryan *et al* used similar methodology to explain coordinate system relationships in [56].)

Relationship	Description	Process
A	Registration cube to linac	Cube is aligned with linac room lasers
B	Tracking markers to registration cube	Rotation and translation parameters applied to markers
C	Tracking markers to world	Transformation calculated from **D** & **F**
D	Markers to camera	Position and orientation of camera determined from markers in camera image
E	3D model to camera	Position and orientation of 3D model in camera image determined by position of camera, from **D**
F	Camera to world	Transformation applied to world coordinates via camera calibration
G	3D model to world	Transformation calculated from **E** & **F**

Table 2.1: Relationships between the individual coordinate systems in Figure 2.14

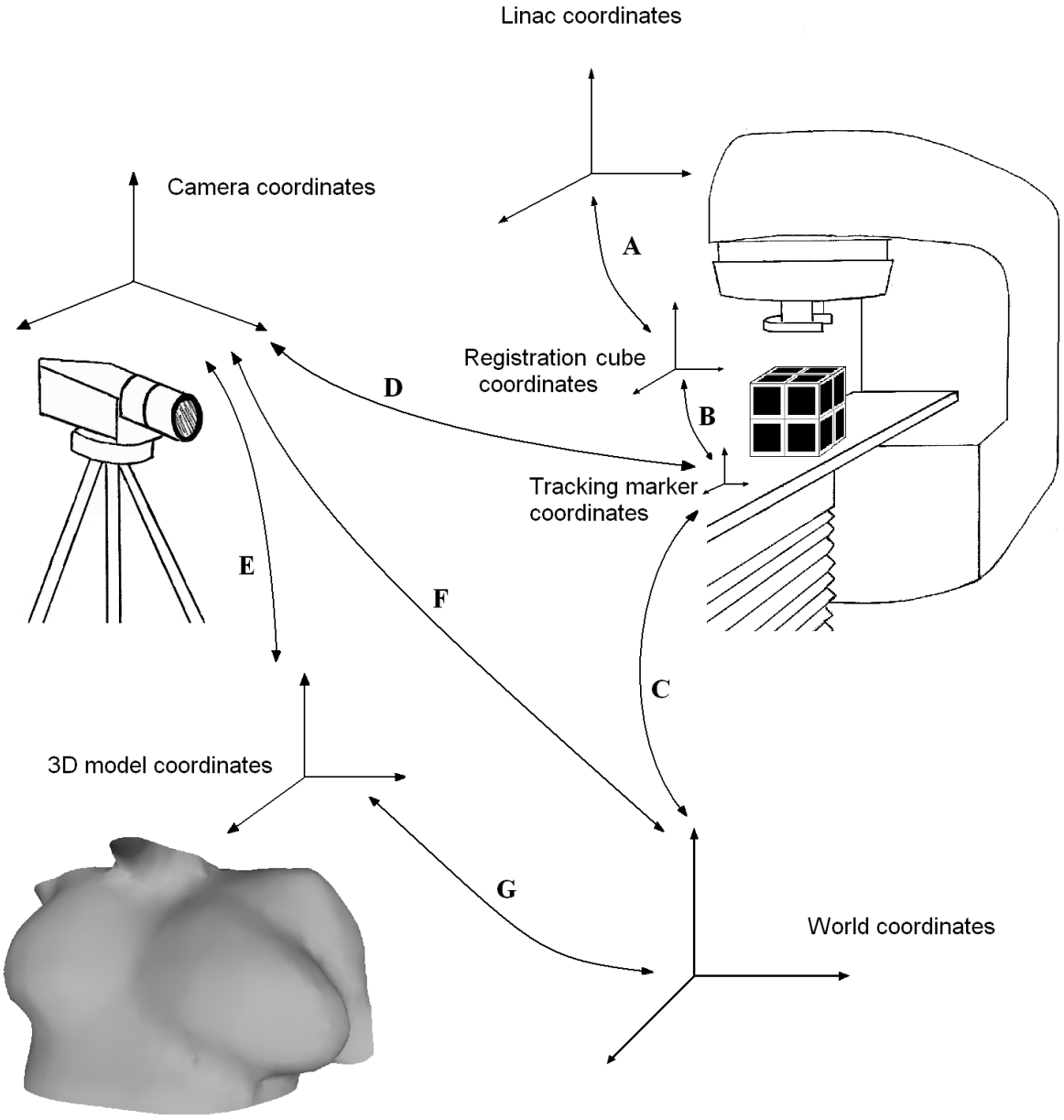


Figure 2.14: Overview of the coordinate systems involved in the AR application. The relationships are explained in Table 2.1

2.8 System Overview

The following steps summarise the operation of the system:

1. Prior to treatment, the patient is admitted for a CT scan to obtain three dimensional data of the anatomy surrounding the tumour.

2. In the 3D CT data, contours are drawn around the tumour, the treatment volume, critical organs and the body outline of the patient. The contour information is stored as a DICOM-RS object.

3. A MATLAB script is used to extract the surface contour from the DICOM-RS object, and 3D models of the patient surface are generated.

4. Surface reconstruction of the pointset model is performed using Meshlab.

5. Prior to treatment, the radiation therapist places the registration cube on the linac treatment couch, and aligns it using room lasers or CBCT, so that the centre of the cube is at the linac isocentre.

6. The camera obtains a live view of the linac and detects the fiducial tracking markers on the registration cube. The 3D models are displayed in the view so that they align with the linac isocentre.

7. The radiation therapist locks the coordinates by pressing "enter." The system no longer tracks the markers on the registration cube.

8. The registration cube is removed from the couch and patient set-up can begin. The radiation therapist can refer to the augmented monitor to align the patient with the correctly positioned contour on-screen.

Chapter 3

Visualisation

During the development of AR programs, it can be tempting to focus solely on technical issues, such as precise coordinate registration and calibration. In this case, these matters were unquestionably crucial; the importance of accuracy in radiotherapy patient set-up is strongly emphasised. However, these efforts would have been wasted if the augmented image were unclear or confusing, so that the actual alignment process became difficult or impossible. An in-depth discussion of the importance of user interaction in AR is provided by Stoyanov *et al* [57].

It was essential to consider how to best relay the computer-generated information to the user, making its operation intuitive and natural. This required qualitative visualisation testing, for which the system had to be set up in either a clinical environment, or in a laboratory set-up that simulated a clinical scenario. The latter method, for the purposes of testing the subjective elements of the system, was a lot more practical. Radiotherapy departments are often busy, and the linacs booked for patients during regular office hours. Performing tests with the system in its early stages of development would have been invasive during a radiotherapy session, and at this stage would have required patient consent and ethical approval.

A laboratory environment was far more convenient, since there were no time constraints to hinder testing. Various system configurations were trialled and evaluated.

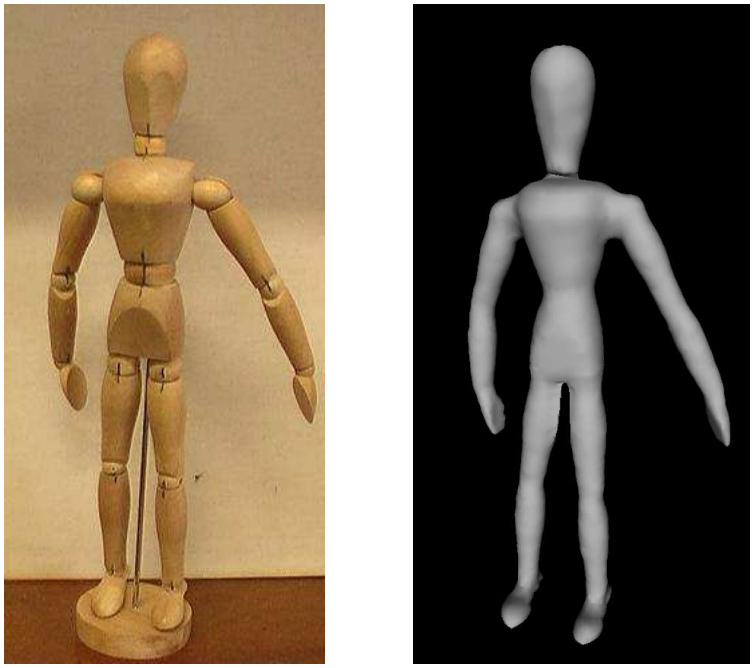
3.1 Test Phantom

In order to assess the system in a laboratory setting, a test model (a phantom) was required to take the place of a real patient. The phantom had to meet two requirements:

- One of the most valuable aspects of this system was that it could allow the radiation therapist to observe deformations to the patient's pose. The phantom therefore required moveable joints, so that deformations could be applied to it.

- The phantom had to be small, so that it would be easy enough to manage in multiple, sequential tests. A large, heavy phantom would have been cumbersome and difficult to reposition time and again in a small lab. Even though a large anthropomorphic phantom would more accurately replicate the geometry of a human patient, this was only significant when evaluating the geometric accuracy of the system (see chapter 6).

A 30 cm tall wooden statue was chosen as the phantom (Figure 3.1). These statues are designed to assist artists in drawing the human figure. As such, they emulate the human form, which made them ideal for use in this project.



(a) The wooden phantom.

(b) Surface reconstructed 3D model.

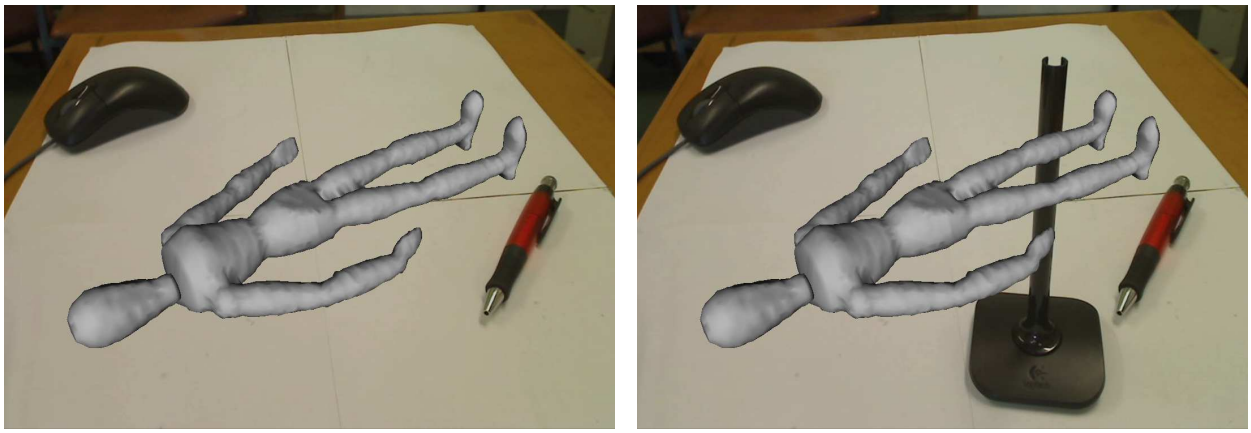
Figure 3.1: The phantom used in visualisation tests.

3D models of the phantom had to be generated for use with the AR software, and these were obtained by performing a CT scan on it, as a radiotherapy patient would receive. 142 transverse slices were obtained, of 2.5 mm thickness. A MATLAB software package for radiotherapy treatment planning research called CERR [58] was used to generate the surface contour of the phantom and store it as a DICOM-RS object. The 3D models were exported as described in section 2.1.

3.2 3D Model Representation

Augmented reality applications seek to fuse virtual objects with the real world, so that they appear to exist in real-space. Spatial registration of the real world coordinates with the coordinates of the 3D model allow the correct geometric transformations to be applied so that it is displayed in the appropriate size, position and orientation. However, its apparent distance from the camera is only an illusion created by the geometric transformations. It does not literally interact with real objects in the scene.

To demonstrate this, Figure 3.2a shows the 3D model of the virtual phantom overlaid on a camera image. The way in which the model is positioned gives the illusion that it exists in the real world and is lying on the bench. However, when an object is placed in the foreground, visual occlusion does not occur as expected (Figure 3.2b). The illusion of the phantom lying on the bench is ruined; it now merely appears as an image of the phantom



(a) The virtual phantom is overlaid on an image of a workbench, appearing to lie down on the surface.

(b) An object is placed in the foreground, ruining the perceptual effect.

Figure 3.2: Visual occlusion problem.

superimposed over an image of the bench.

A similar effect occurred when trying to bring the real phantom into alignment with the 3D model. The model ceased to appear as an object that existed in real-space and seemed more like a superimposition in the foreground. The desired effect was to have the real phantom and the virtual phantom appear to exist in the same space, so that aligning the two became intuitive. Ideally, someone who had no experience with the system should have been able to perform this alignment with ease.

This problem falls under the category of what is known as photometric registration [59]. Photometric registration is the process of correctly representing an object in terms of matters such as image quality, depth perception, occlusion, lighting and shading. To create the correct perceptual illusion, a number of visualisation techniques were applied to the representation of the virtual phantom.

3.2.1 Transparency

When the virtual phantom was displayed as a solid, opaque model, visual alignment was extremely difficult. When the real phantom was aligned correctly, it was completely obscured by the virtual overlay. In this situation, there was no way of telling if the real object were in spatial alignment with the virtual one, or if it were behind it, at some point in the background (Figure 3.3).

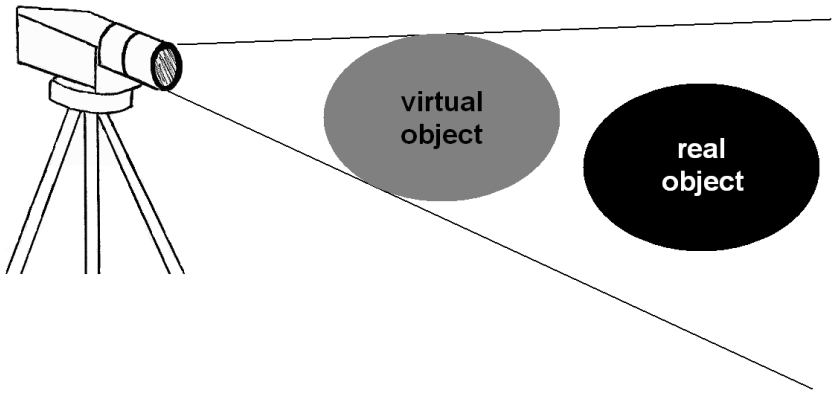
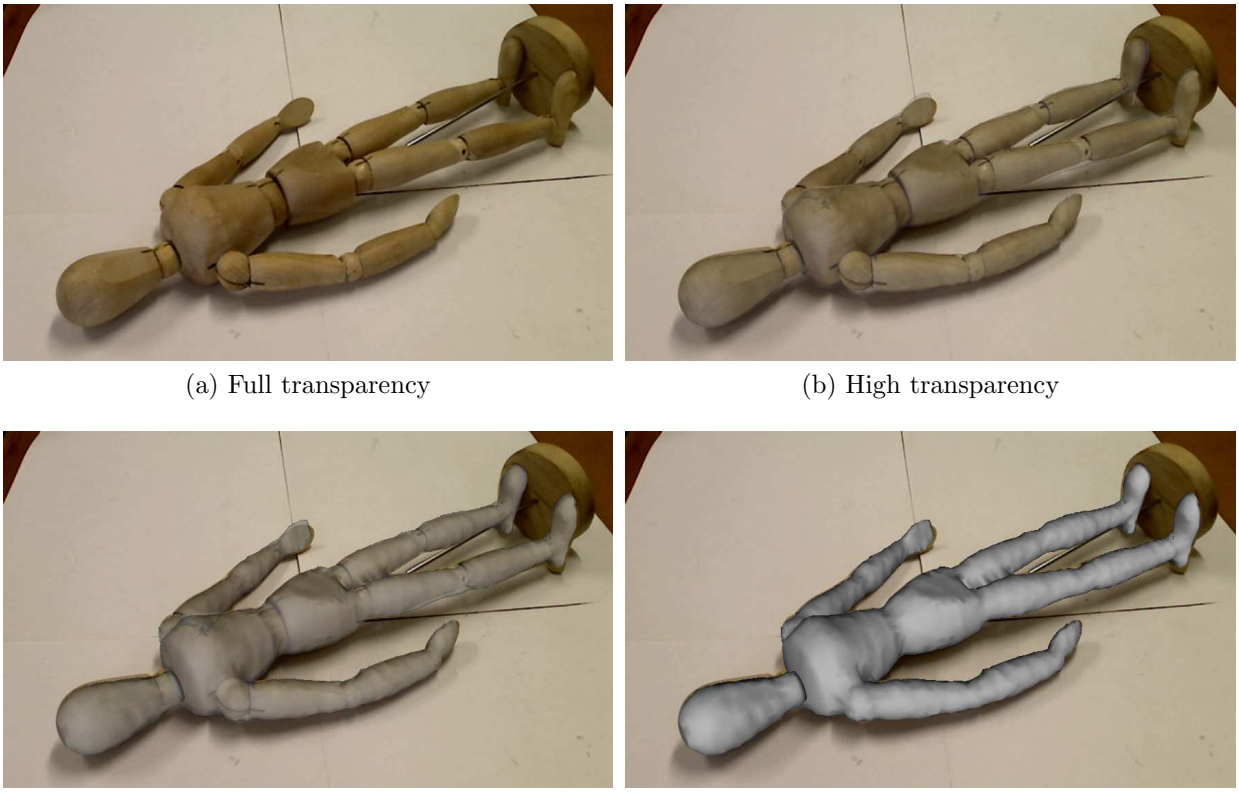


Figure 3.3: From the camera's point of view, the real object is in line with the virtual object. Since the virtual object is opaque, the user cannot tell that the real object is situated behind it, and that they are not in spatial alignment.

The system was intended to allow non-rigid deformations to patient pose to become visible. With an opaque model, even if the patient were aligned with the linac isocentre, it would still be possible for deformations to be obscured from sight. An option to make the model partially transparent was essential. The user could then visualise both the 3D model and the real patient simultaneously, and know that they were spatially aligned when the outlines of both coincided (Figure 3.4). This function was added with the option of adjusting the level of transparency, so that the model could be reduced to near, or complete transparency when the user required a clearer view.



(a) Full transparency

(b) High transparency

(c) Low transparency

(d) Opaque

Figure 3.4: Varying transparency.

3.2.2 Wireframe

Displaying a translucent augmentation allowed a clear fusion of the real world and virtual objects, but did not provide the user with a true sense of depth. The semi-transparent augmentation was still only a visual superimposition over the live image and did not provide much detail of the surface topography of the object. Even though the user may have been able to align the patient to the outline of the virtual object, the effect may have been misleading, and with no perceived sense of depth it would have still been possible to misalign the patient, as in Figure 3.5.

To resolve this issue, an option to combine the transparent model with a wireframe representation of the virtual object was included (the acquisition of which was described in section 2.1). To facilitate a better visual identification of the curvature of the surface, the wireframe was comprised of discrete closed contour lines around the surface of the virtual model. To avoid an overload of information, the contour lines were only in the transverse plane (Figure 3.6). The obvious directionality allowed surface topography to be visualised easily.

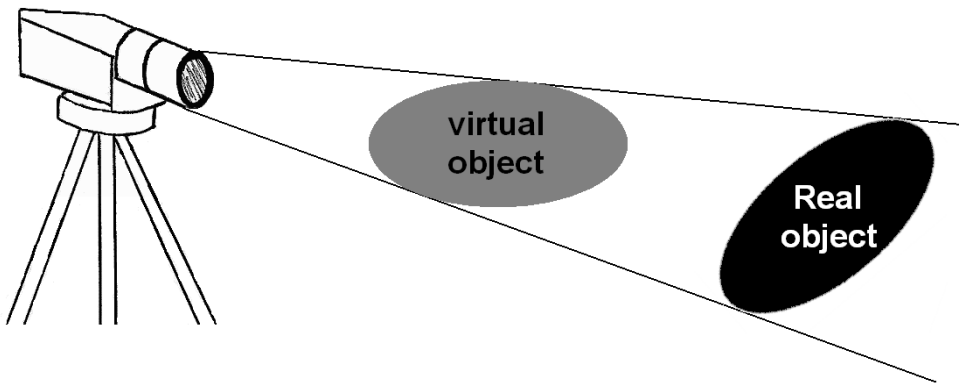


Figure 3.5: From the point of view of the camera, the outlines of the virtual and real objects will appear to coincide, even though the two objects are not in spatial alignment.

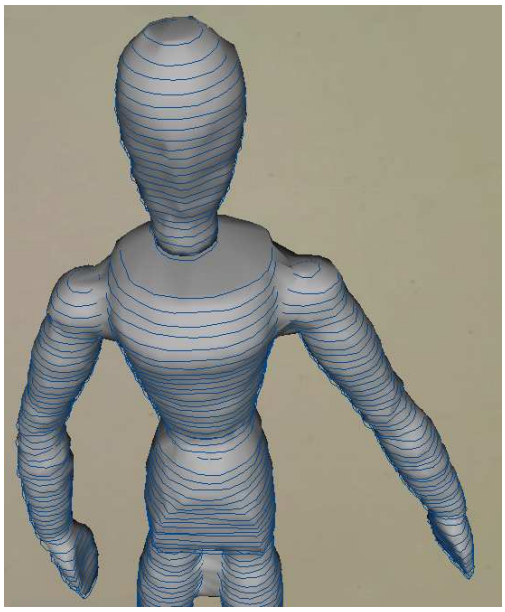


Figure 3.6: When displayed with a wireframe around it, the topography of the model becomes distinguishable.

3.2.3 Backface Culling

When transparency of the 3D model was enabled, the user was able to see both the front and back surfaces, as in Figure 3.7. This convoluted the image, making the model difficult to see through. A graphical technique called "backface culling" was implemented to resolve this issue.

Backface culling removes any faces on the reverse side of the model from the visualisation. Each face in the 3D model had a normal associated with it, which pointed away from the face in the direction that it was oriented. Backface culling removed the faces that have normals pointing away from the camera's plane of view (Figure 3.8).

In this system, backface culling was able to add the illusion of the virtual object occupying the same space as the real one. The virtual phantom then appeared to wrap around the surface of the real phantom, rather than just float above it as a visual superimposition (Figure 3.9a). This effect would be especially useful in situations where the CT scan does not encompass the whole of the patient's body (as is usually the case). In such a case, the upper and lower bounds of the 3D model would be open, revealing the hollow interior of the model. Backface culling makes the interior invisible (Figure 3.9b).

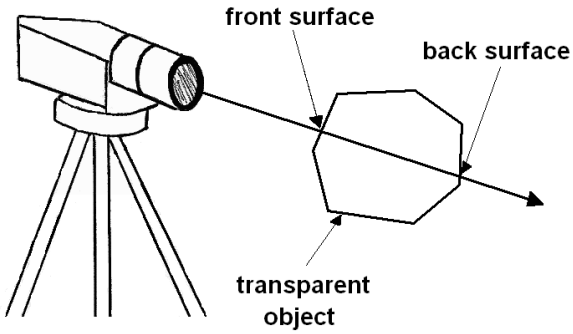
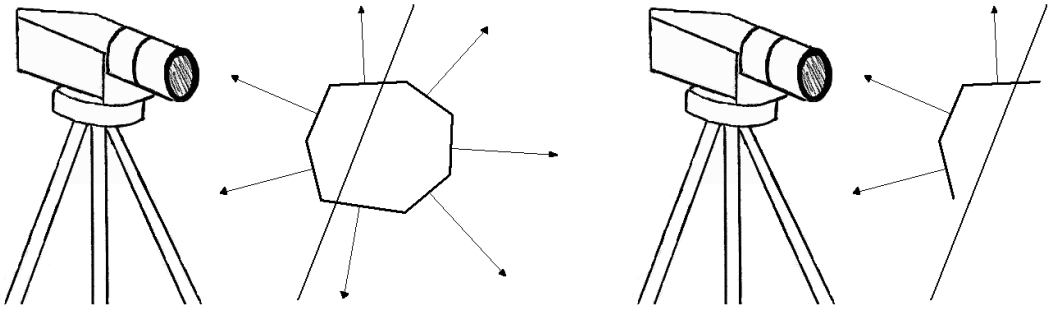


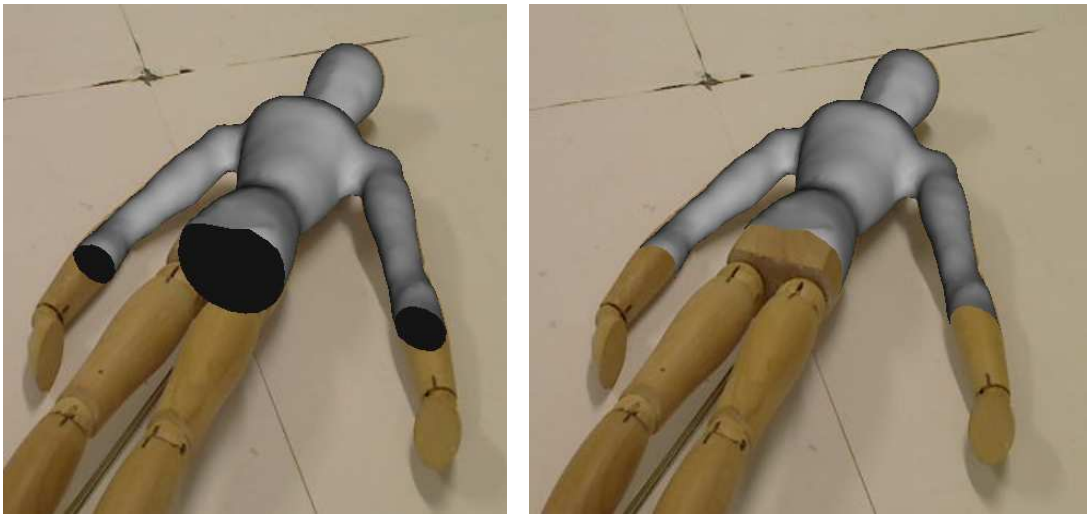
Figure 3.7: The camera observes both the front and back surfaces of the transparent virtual object.



(a) The faces on the virtual object have normals that describe their orientation.

(b) Those that have normals that face away from the camera's plane of view are removed, so that only the front of the object is visible.

Figure 3.8: The process of backface culling.



(a) No backface culling. The hollow interior of the 3D model is visible.

(b) When backface culling is enabled, the virtual phantom appears to wrap around the surface of the real phantom, giving the proper illusion that the two are aligned.

Figure 3.9: Backface culling applied to the phantom representation.

3.2.4 Position Smoothing

For all the pose-estimation precision of the professional edition of ARToolKit, there was still a degree of error involved in tracking the markers on the registration cube. As the cube was moved further from the camera, the software had difficulties resolving the tracking markers in the camera image, due to the finite resolution. The precision of the transformations that described the camera's position suffered and the stability of the augmentation decreased (this effect is explained in greater detail in section 5.2). This resulted in the 3D model "wobbling" as the software updated the position transformations from frame to frame, even when the camera was stationary.

To counter this effect, rather than updating the transformations in every frame of the video input, the software was made to "smooth" the transformations. This process reduced the differences in the translation and rotation matrices between subsequent frames to 15% of their magnitude. This meant that the magnitude of the wobbling effect was reduced to 15%, resulting in a very stable pose estimation that had little variation from one frame to another.

The downside to this smoothing method was that if there were sudden movements to the tracking markers, the virtual object would lag for about a second, before catching up with the movement. In a dynamic environment where the tracking markers and the camera constantly move about, a smoothing feature would be unsuitable. However, in this system the camera was fixed in place and, once registration had been performed with the cube, the fiducial markers were no longer tracked.

Chapter 4

Layout Design

The physical arrangement of the equipment had a significant impact on the system in terms of its usability and accuracy. The placement of the camera and monitor not only had an effect on the generation of geometric transformations for pose-estimation, but also on the physical and cognitive efforts required by the user to operate it. This chapter discusses how the equipment could be physically arranged to produce the best results.

4.1 Camera Placement

The position and angle of the camera had to be considered carefully. Alignment had to be performed in three dimensions (the xy plane was the horizontal plane of the treatment couch and the z direction referred to the couch height). The camera therefore had to be placed at an angle that would allow the accuracy of the patient's position in all three directions to be clearly visualised.

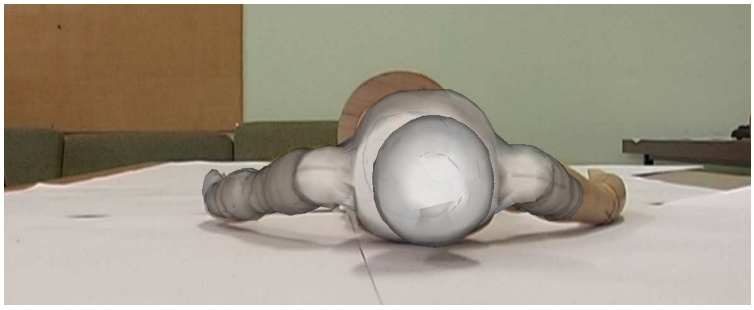
In order to properly observe non-rigid patient deformations, placing the camera at a steep angle above the patient was advantageous. Figure 4.1a shows how such an angle made deformations in the xy plane clearly visible, compared to a low angle. However, information regarding the height of the phantom from the ground was limited. In a clinical setting, raising or lowering the treatment couch to the correct level would have been difficult.

Conversely, with the camera positioned so that it was level with the couch, viewing the

patient from the side, it was easy to see if the phantom was positioned correctly in the yz plane, and couch-height errors become apparent (Figure 4.1b). Deformations, though, became less obvious. Camera angles along the y axis (i.e. looking directly at the patient's feet or head) were avoided, as most of the body became obscured (Figure 4.1c).

(a) Coronal view. The right arm is clearly misaligned.

(b) Sagittal view. It is obvious that the phantom is not at the correct height.



(c) Transverse view. Visual information is limited.

Figure 4.1: Viewing the phantom from various camera angles.

Apart from the visual advantages and disadvantages of positioning the camera in certain orientations, the limitations of the ARToolKit tracking software had to be considered. When the square tracking markers were visible, the software detected the corners of each of the markers, and used these vertices as reference points when it estimated the pose of the 3D model. If the camera were oriented so that only one cube face was visible, then all of the reference points would have been in a single plane (Figure 4.2). Virtual objects that were further from this plane would not be displayed with the same degree of precision as those that were close to it.

The best results could be obtained when the camera was positioned so that, during the registration process, as many cube faces as possible were visible (i.e. three). Taking this into account, as well as the visual advantages associated with certain camera angles, the most advantageous camera position was diagonally adjacent to the treatment couch, elevated approximately 45° above it, as in Figure 4.3.

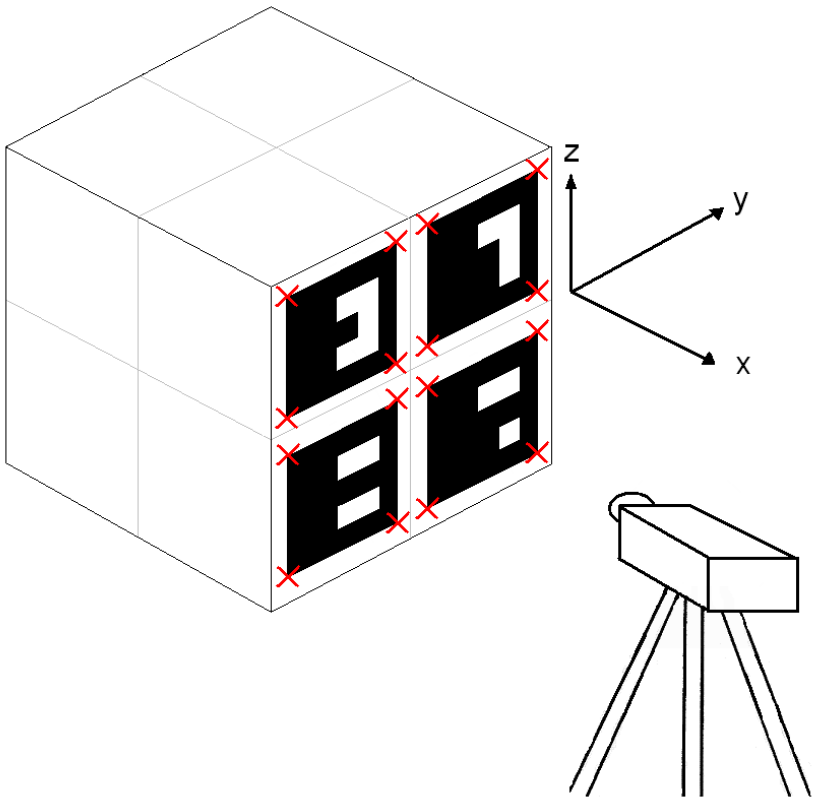
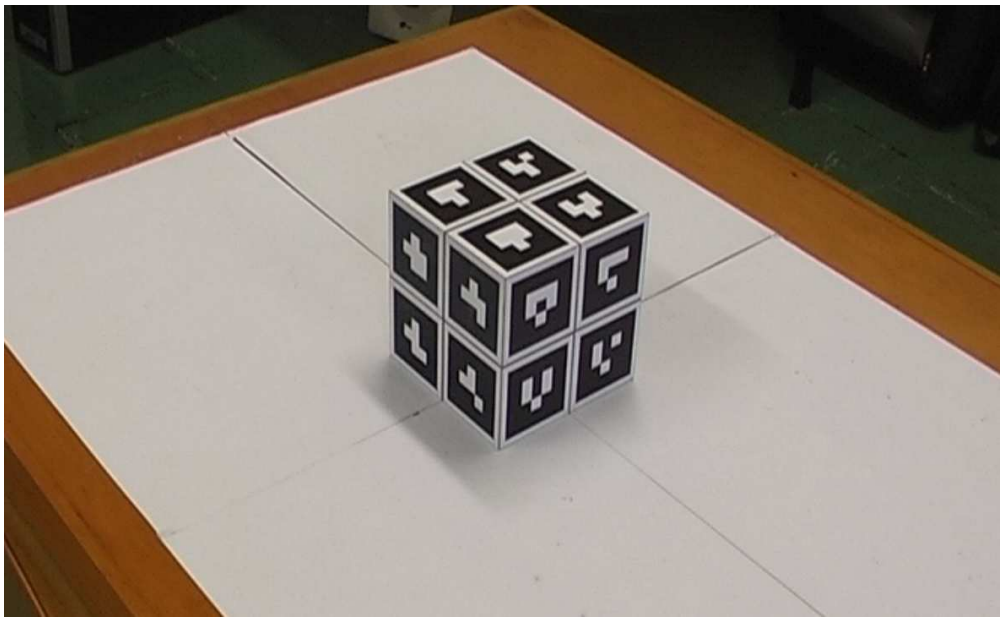
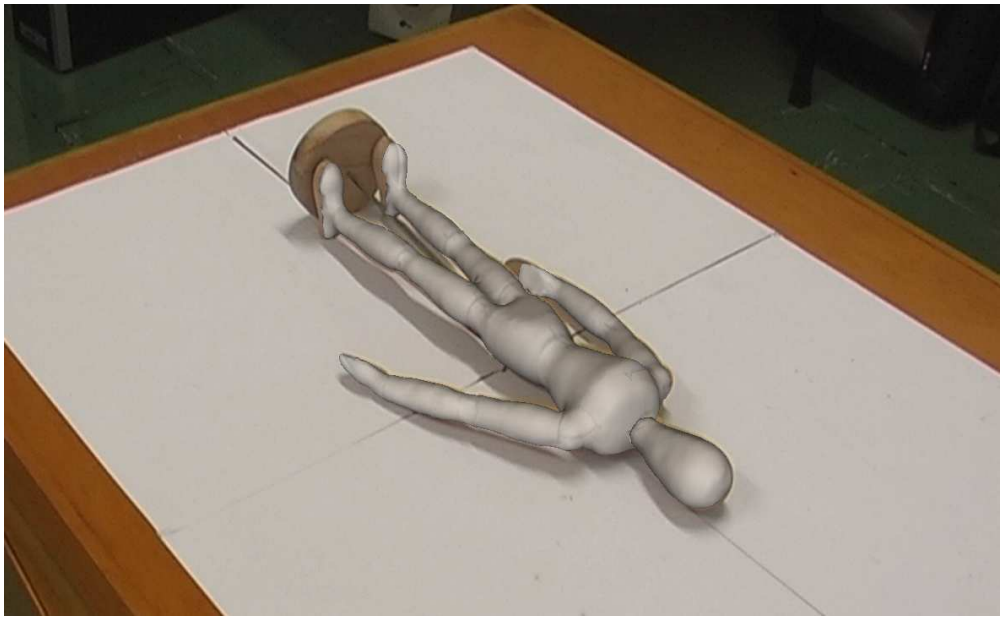


Figure 4.2: When only one cube face is visible to the camera, all the points of reference (red crosses) are within a single plane (the yz plane). Virtual objects that are distant to this plane would be positioned with less precision.



(a) The camera is placed at one of the corners of the workbench and elevated so that three of the cube faces are visible.



(b) From this camera angle, it is possible to align the phantom confidently, while maintaining the ability to observe non-rigid deformations.

Figure 4.3: Ideal camera angle.

Since this system was supposed to invade on the radiotherapy procedure as little as possible, the logistics of the camera set-up had to be considered carefully too. Although an augmented monitor system did not involve cumbersome headgear or handheld equipment, it still required equipment that occupied additional space. In a radiotherapy clinic, the camera should be placed where it does not impinge on the space of the medical staff and, since it must remain perfectly still throughout treatment, where it is unlikely to be bumped.

4.2 Multiple Camera Approach

When a single camera was used for video acquisition, the image that the user saw in the monitor was two-dimensional, whereas the patient had to be aligned in three dimensions. Even when the camera position was optimised as above, and the visual effects that were outlined in section 3.2 were applied, a true sense of depth was lacking, and alignment errors could easily occur.

If the camera could have been shifted after initial registration with the cube, the phantom could have been aligned with the virtual object as closely as possible, the camera shifted and the process repeated as many times as necessary. However, the camera had to remain stationary, and even if tracking were possible, repeatedly shifting the camera and performing realignment would have been time-consuming and awkward.

Using two cameras for video acquisition solved this problem. A two-camera system provided stereoscopic vision, so that between them they covered all three spatial dimensions and allowed the user to position the phantom properly. This is illustrated in Figure 4.4. In the left camera image, the phantom is clearly misaligned. In the right image though, there is no visible error. If only the right camera image were used, it would be easy to mistake the phantom for being perfectly aligned.

With the two-camera system, optimal results could be obtained when the first camera was elevated above the patient, allowing deformations to be visualised clearly, and the second was level with the treatment couch, so that it showed the couch height.

Note that having two AR programs running (one for each camera) placed significant strain

Figure 4.4: Simultaneous video acquisition from two cameras.

on the CPU and GPU. An alternative configuration was the use of two computers, one to process the input from each camera, and the images displayed on two separate monitors. This would, however, incur additional cost to install, and would occupy more space within the linac bunker.

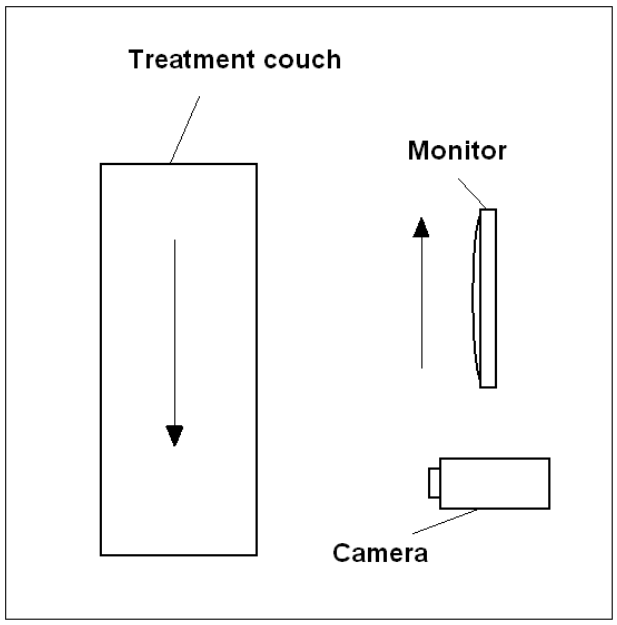
In a similar way, more than two cameras and monitors could be installed to obtain views from any number of angles. In principle, each camera would provide the user with more spatial information. However, the cognitive effort required to mentally process the images on all the monitors would increase, resulting in an overload of information. The monetary cost would also accumulate with each additional camera, monitor and computer, as would the space they occupy. It was therefore decided that two cameras were sufficient.

4.3 Monitor Placement

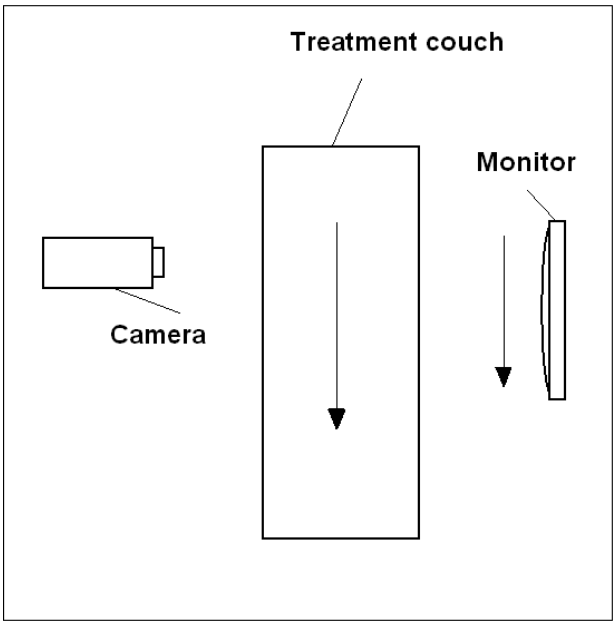
The usability of the system was affected by the placement of the monitor as much as it was by the video cameras. As with the cameras, the monitor had to be positioned where it was unlikely to create a physical obstruction. In a clinical setting, mounting it to a wall would have been appropriate, but the size of the linac bunker would have to be considered. Typically there is a lot of space between the treatment couch and the walls, so that movement within the room is not hampered. The monitor would therefore be distant to the user and difficult to view.

The relative location of the monitor to the camera also had to be considered. If the monitor and camera were adjacent to one another and facing the same direction, movements in the real environment were reversed in the monitor (Figure 4.5a). This sort of arrangement was confusing and unintuitive. When presented with a dynamic image that reproduces the environment around it, the user expects movements in the image to correspond to those of the real world.

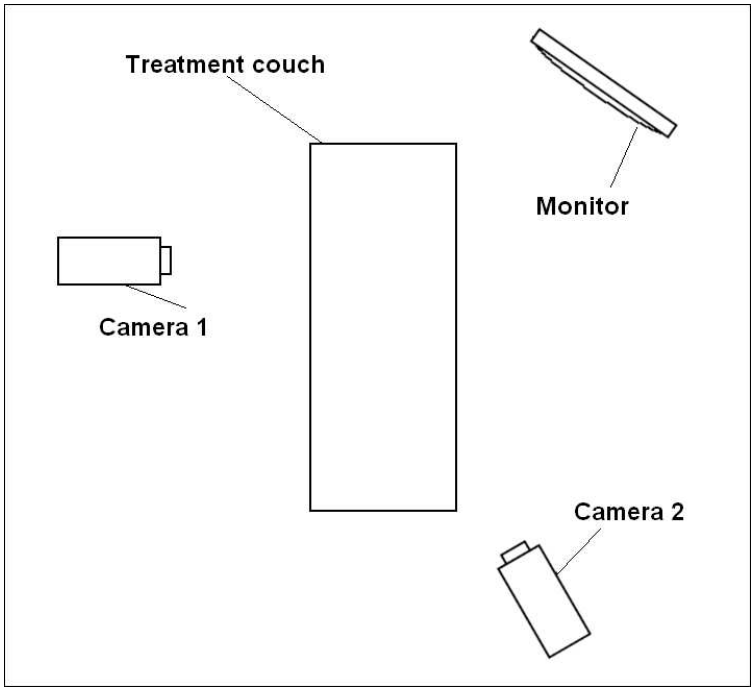
It was found that it was best to position the monitor opposite the camera, as in Figure 4.5b. This resulted in a far more natural correlation between the real world and the augmented view on the monitor. Similarly, when two cameras were in use, the best results could be obtained if the monitor opposed both cameras, when possible (Figure 4.5c).



(a) When both the camera and monitor face the same direction, movement on the treatment couch occurs in the opposite direction on the monitor.



(b) When the camera and monitor face opposite directions, the direction of movement on the couch corresponds to that on the monitor.



(c) When two cameras are used, the monitor should still be opposite both of them.

Figure 4.5: Monitor placement relative to the camera.

Chapter 5

System Accuracy

An ideal patient set-up system in radiotherapy would allow the patient to be positioned on the linac exactly as they were in the planning CT, so that the treatment isocentre aligns exactly with the linac isocentre, and the patient's pose is identical. In reality though, an ideal system that meets these requirements is not possible. Regardless of the set-up method, a small degree of error is unavoidable.

In this system, the potential sources of set-up error were identified individually so that efforts could be made to minimise them. The individual sources of error were split into three categories:

- Machine errors, which were attributable to the limitations of the system hardware and software.
- Human errors when manually aligning the patient.
- Variations in patient anatomy.

The latter source of error cannot be regulated. *Intra-fraction* variations (such as respiration and patient motion) and *inter-fraction* variations (such as weight loss) are inevitable over the course of radiotherapy regimes. Evaluation of the system with real patients was outside the scope of this project, so this source of error was ignored. It is worth noting, however, that if it were implemented clinically, one of the benefits of the system is that it would allow these variations to be observed.

In this chapter, the sources of machine error are critically identified, and the measures that were taken to minimise them are described.

5.1 Registration Cube Geometry

The registration cube was constructed to precise dimensions, with side lengths of 92 mm. These lengths were confirmed via measurements with vernier calipers, and the uncertainty in each was deemed to be no greater than 0.3 mm. The actual shape of the cube was also considered, since any warping would have shifted the tracking markers attached to the faces, creating a discrepancy between their actual locations and the position parameters that were entered into the system (section 2.6.1). For this reason it was constructed of perspex, a robust material that is resistant to warping, so that the faces would remain completely orthogonal. It was also made entirely solid, with no hollow interior.

Each of the cube faces (except for the underside) had four tracking markers attached to it. Great care was taken to position the markers squarely, so that their positions correlated exactly with the parameters in section 2.6.1 (see Figure 2.10). The uncertainty in the positions of the markers was estimated to be 0.5 mm.

Registration cube geometry errors made an insignificant contribution to the overall error of the system. With the camera (or cameras) at a distance of, say, one metre (at least a metre would be necessary to encompass the patient within the image), a tracking marker offset of 0.5 mm would be lost to the finite resolution of the camera. At this distance, tracking errors would be dominant. Errors in the geometry of the cube and the arrangement of tracking markers were taken to be negligible.

5.2 Tracking Accuracy

The accuracy of the tracking software was one of the greatest sources of error in the system. Despite the technical advances of augmented reality tracking technology over the last few years, AR is still a relatively young field and has limitations. The accuracy with which

the system determined the position of the camera (and hence the geometric transformations which governed pose-estimation of the 3D model) depended on how accurately it could detect the tracking markers. This was subject to:

- Room lighting.
- Camera focus.
- Camera resolution.
- Camera angle.
- The distance of the camera from the markers.
- The size of the tracking markers.

Room lighting dependency and camera focus issues were not difficult to overcome. The brightness, contrast and focus in the camera feed were simply adjusted so that there was a sharp delineation between the black and white of the tracking markers.

The finite camera resolution (2.3 megapixels) was far more problematic. The system calculated the position of the camera by detecting the markers and using the four corners of each marker as geometric reference points. The further the camera was from the markers, the less precisely these corners could be resolved.

The problem worsened when displaying 3D objects that were much larger than the size of the cube. Since the objects were essentially displayed with respect to the reference points (the corners of the markers), the system had to extrapolate further when calculating transformations for large objects, resulting in greater potential for error.

This effect is illustrated in Figure 5.1. When two concentric circles are displayed relative to the registration cube, further extrapolation is required for the pose estimation of the larger of the circles. The magnitude of the pose-estimation errors are **a** and **b** for the small and large circles respectively.

Note that the nature of this error was rotational, rather than translational. However, the magnitude of the uncertainty due to this error was difficult to measure directly, and so was estimated to be 0.5° in roll, pitch and yaw.

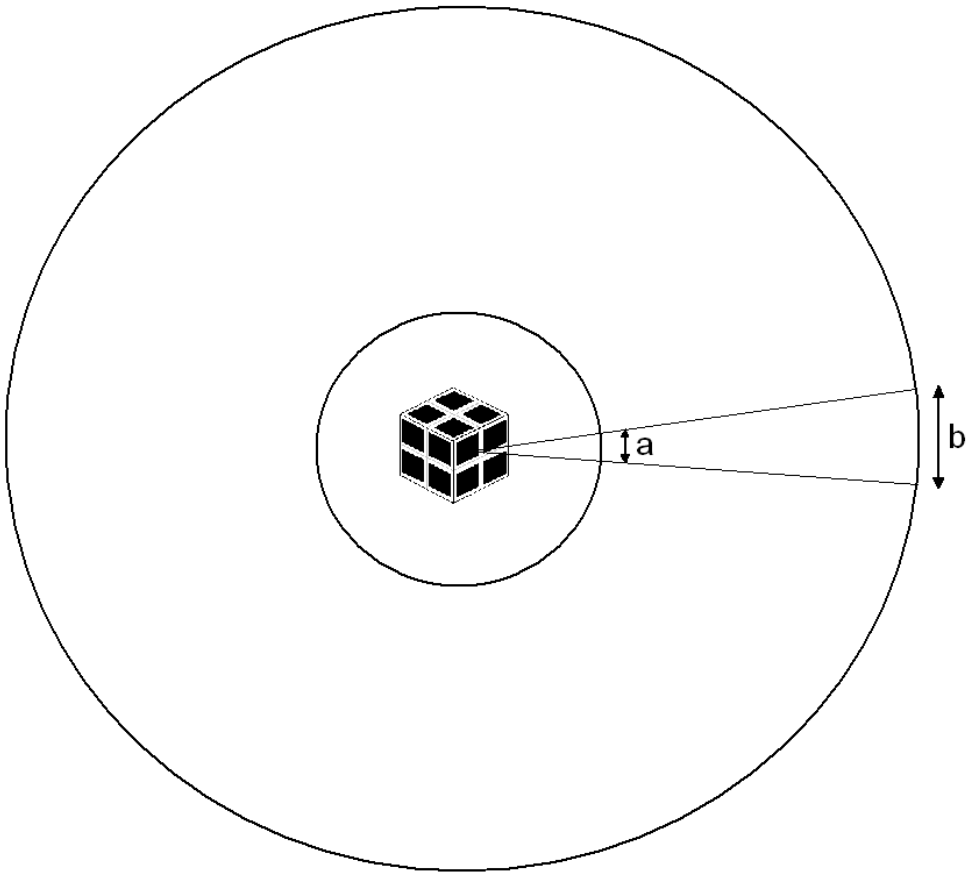


Figure 5.1: Extrapolation in pose-estimation for small and large objects.

5.3 Camera Calibration Accuracy

The application that was used to calibrate the camera to compensate for distortion was discussed in section 2.6.2. After calibration had been performed, however, a form of measurement was required to determine how successful the process had been. A simple visual method was devised.

A wireframe model of a simple cube was generated, which had the same dimensions as the registration cube. This model was loaded into the AR system and displayed so that its centre matched that of the registration cube. When the sides and corners of the virtual cube lined up with those of the real one, it confirmed that camera calibration had been successful (Figure 5.2).

The calibration program produced varied results. Sometimes there would be an obvious discrepancy between the real and virtual cubes, other times they would align with no discernable offset. It seemed that the best results could be obtained if the following steps were followed when obtaining snapshots of the calibration pattern:

1. The calibration pattern was held at the same distance from the camera as the registration cube.

2. Three to four snapshots were obtained at this distance, with the pattern held at a different angle each time.

3. Steps 1 and 2 were repeated twice more with the pattern held closer to the camera each time.

It was decided that every time camera calibration was carried out, the test using the cube wireframe should be performed to check that the calibration had produced ideal results.

The digital video cameras that were used in this project produced a prominent barrel distortion effect that was particularly noticeable at the edge of the video image. Even when the calibration process gave optimal results, the intentional distortion that was imposed on the 3D model did not perfectly match the camera distortion towards the edge of the image. Figure 5.3 demonstrates this effect. At the upper-left edge of the cube, the wireframe does not line up perfectly with the real cube. To minimise the impact of this effect, it was best to ensure the cube was positioned toward the centre of the video image, where little warping to the tracking markers occurred.

This error was not constant; it varied between the centre of the image and the edge. Again, the uncertainty attributable to this error was difficult to measure, so it was estimated to be (on average, between the centre of the image and the edge) a 1 mm translational uncertainty.

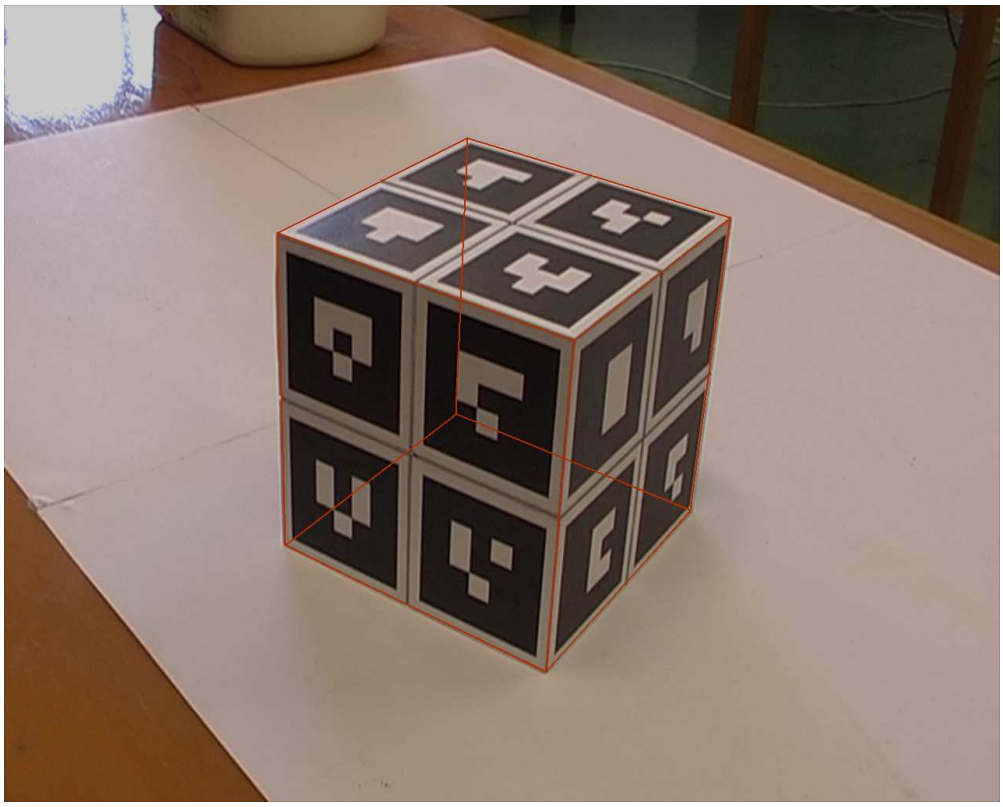


Figure 5.2: Cube wireframe (orange) aligned with the real cube.

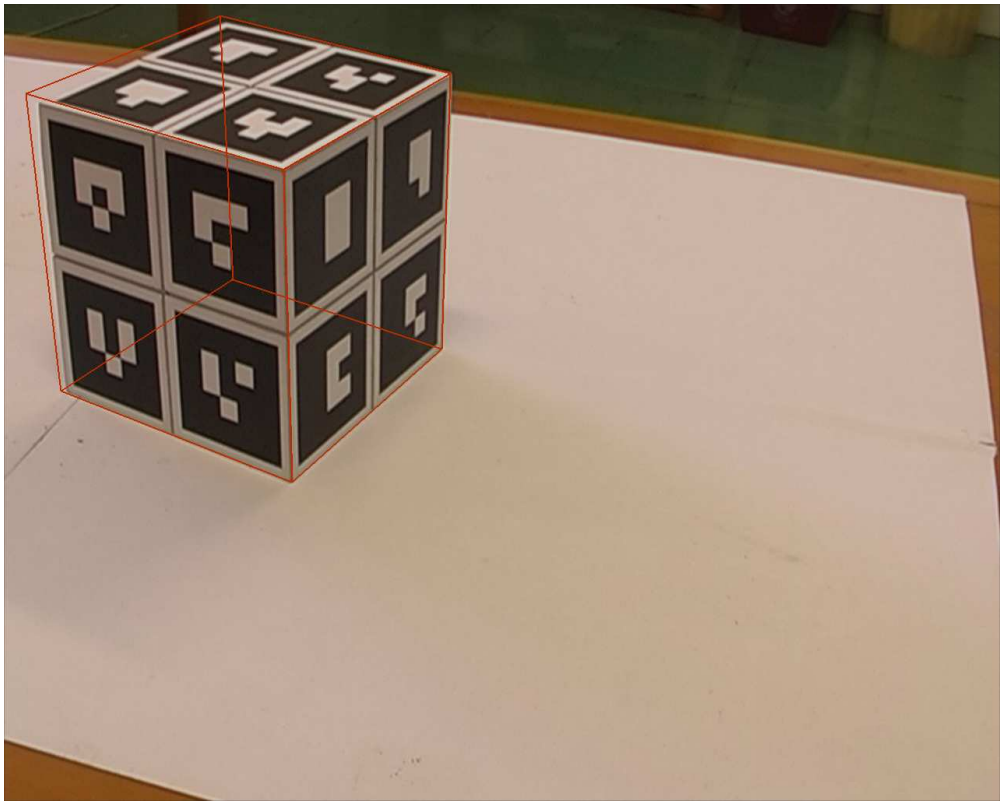


Figure 5.3: Barrel distortion effect on model superimposition.

5.4 Room Laser Accuracy

The accuracy with which the registration cube could be aligned with the linac isocentre depended on how close the intersection point of the lasers was to the isocentre. If the lasers were slightly off, then the centre of the cube would not correspond to the linac isocentre, and the error would propagate to patient set-up. Even if the intersection point correlated accurately with the isocentre, however, errors could still occur if the laser planes were not entirely orthogonal.

Since the isocentre is a virtual point in space, there was no obvious way of determining the proximity of the laser intersection and the isocentre. The American Association of Physicists in Medicine (AAPM) recommends daily tests to ensure the intersection of room lasers is within 2 mm of the linac isocentre [60], so uncertainty relating to laser accuracy was taken to be 2 mm.

Alternatively, cone-beam CT could be used to align the registration cube on the treatment couch, as described in section 2.4. CBCT alignment provides a more precise correlation between the treatment isocentre and the linac isocentre, so the uncertainty due to registration cube misalignment could be minimised.

The sources of uncertainty that were discussed in this chapter could be combined into a single figure that represented the entire machine uncertainty. By summing the uncertainties from individual sources, the total translational error in pose-estimation was estimated to be 3 mm, and the rotational error to be 0.5° in all three planes.

Chapter 6

Clinical Experiment

The previous chapter discussed the sources of machine error, due to limitations of the software and equipment, but did not take into account human error. Even if the AR display showed the patient contour aligned and oriented perfectly, the accuracy of patient set-up would still depend on how well the radiation therapists could utilise the system.

To properly gauge the system, it was set up and examined in a linear accelerator bunker. Permission was granted by the Christchurch Hospital Oncology Department to use one of their linacs, a Varian CBCT-equipped unit. The experiment was performed after-hours, so that it would not encroach on time designated for patient treatments.

6.1 Method

Testing the performance of the system required a phantom to be set up on the linac treatment couch, but the small phantom that was described in section 3.1 was not suitable at this stage. A larger anthropomorphic phantom which could replicate not only the figure of a human, but also the size, was required. For this purpose, the hospital permitted the use of a RANDO ® phantom [61] (Figure 6.1). These phantoms are designed for dose distribution mapping in radiotherapy, and contain a natural skeleton on the interior. They are divided into 2.5 cm sections which can be removed to insert dosimetry devices.

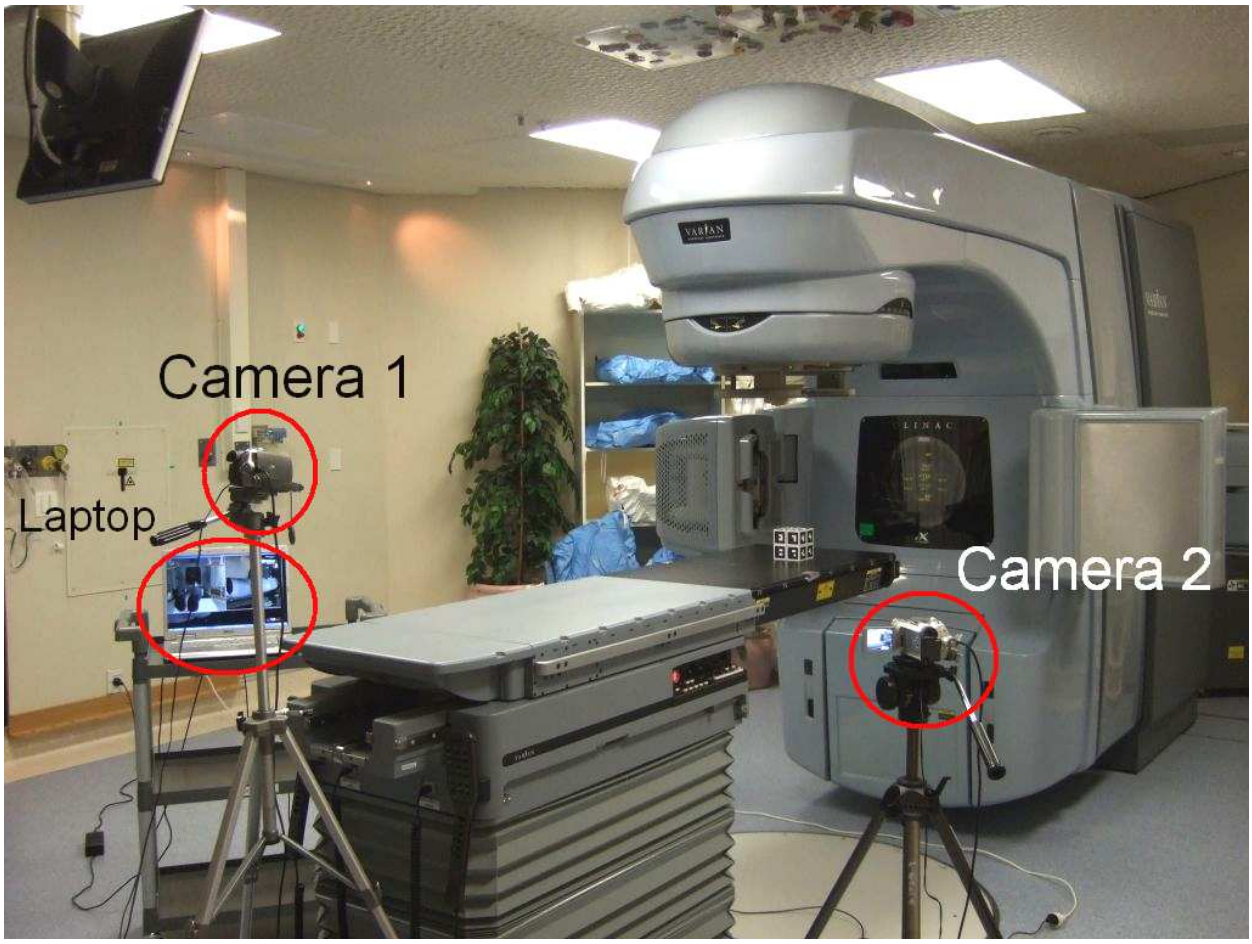
Figure 6.1: The RANDO ® phantom (left) and 3D model (right).

A scan was obtained of the RANDO ® phantom with a GE 64 slice CT scanner at Christchurch Hospital. 385 slices of 2.5 mm thickness were obtained, and the treatment isocentre was positioned in its head. A 3D model was obtained using the MATLAB script described in section 2.1.

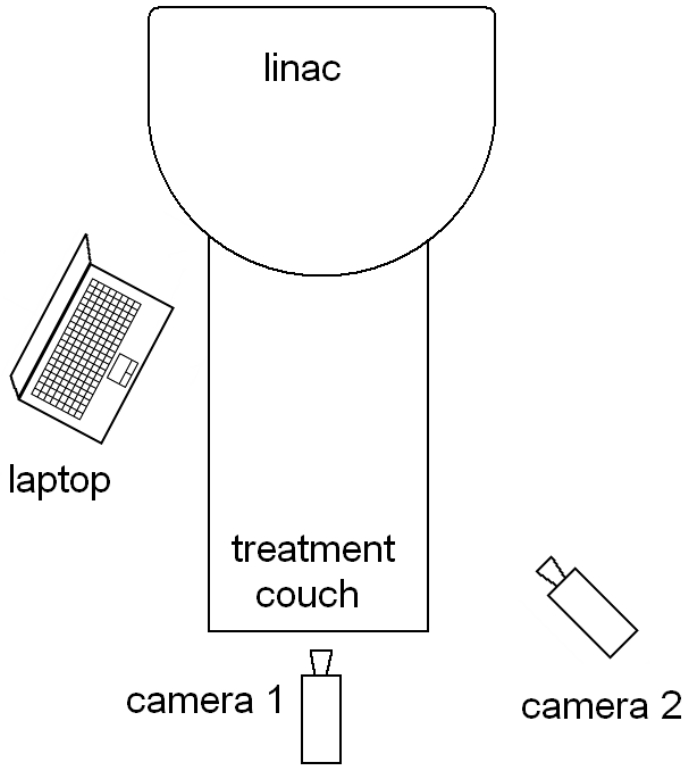
To begin with, the two digital video cameras and laptop were set up in the linac bunker. As was outlined in chapter 4, the layout of the cameras and monitor impacted on the usability of the system, and so was chosen carefully. The layout shown in Figure 6.2 was selected after some experimentation.

Both cameras were attached to tripods to keep them stable. Camera 1 was positioned at the foot of the treatment couch, 235 cm from the linac isocentre and elevated roughly 30° above it. Camera 2 was level with the couch, on the right side of it, 210 cm from the isocentre.

The AR program was run twice, once for each camera, and the two images displayed on the laptop monitor. The registration cube was then placed on the treatment couch, which was raised to the correct height, and was aligned with the room lasers, bringing its centre into coincidence with the linac isocentre. When it was aligned satisfactorily (Figure 6.3), the "enter" button was pressed on the laptop to disable tracking and allow the cube to be removed, thus completing coordinate registration. The more accurate method using CBCT was considered, but could not be implemented due to technical difficulties in transferring the CT data to the linear accelerator treatment planning system.



(a) Photograph of the set-up.



(b) Schematic of the set-up.

Figure 6.2: Overview of the AR system set-up at Christchurch Hospital.

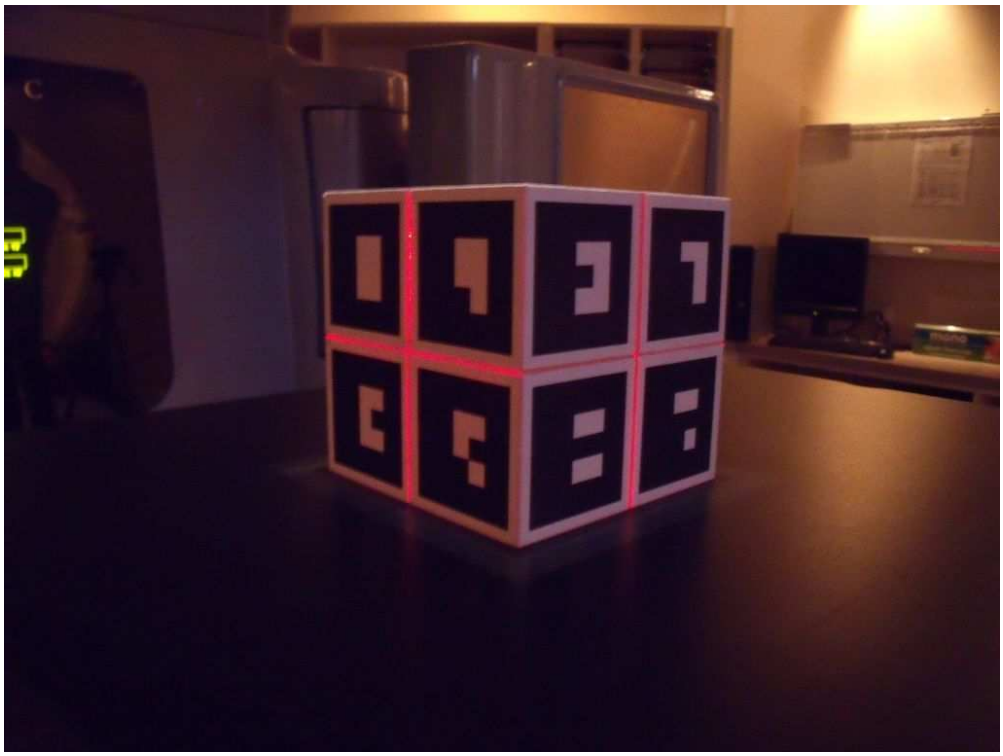
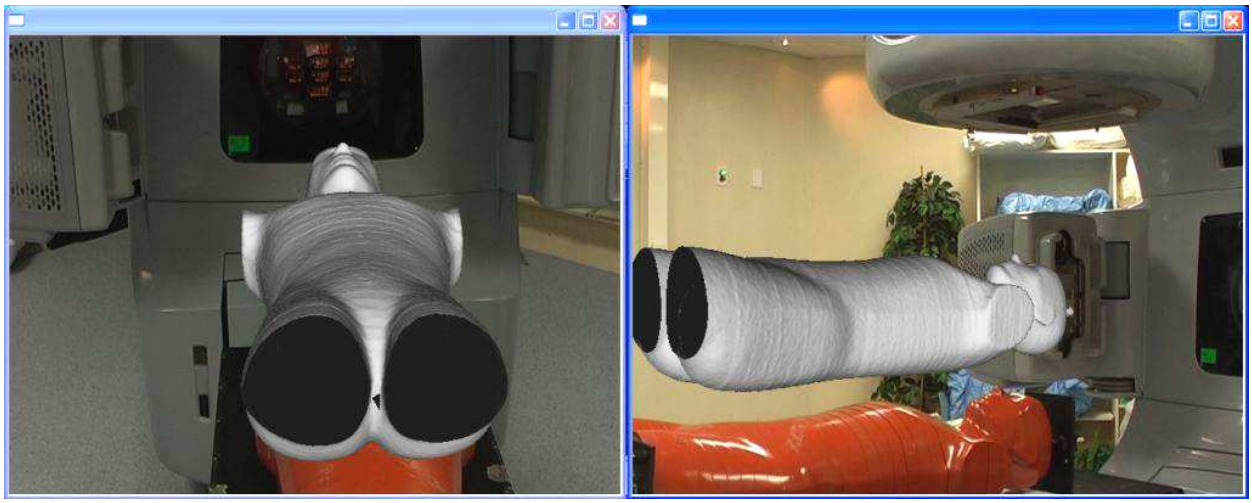


Figure 6.3: The registration cube aligned with room lasers.

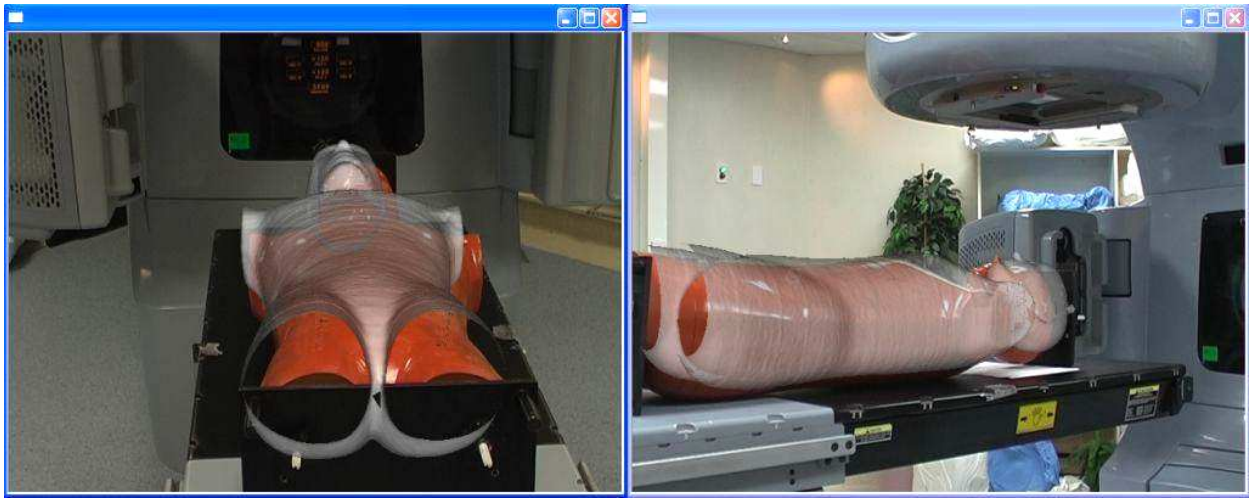
The 3D contour of the phantom was displayed and the real phantom was placed on the treatment couch (Figure 6.4a). The couch height was altered again, until the phantom was at the same height as the 3D model (Figure 6.4b). The phantom was then shifted horizontally to align it with the model (Figure 6.4c).

This test confirmed that visual alignment with the system was intuitive and straightforward. However, it did not give any immediate indication of how accurate the set-up had been. In other words, there was no way of telling how close the treatment isocentre of the phantom was to the linac isocentre, or how accurately it had been oriented.

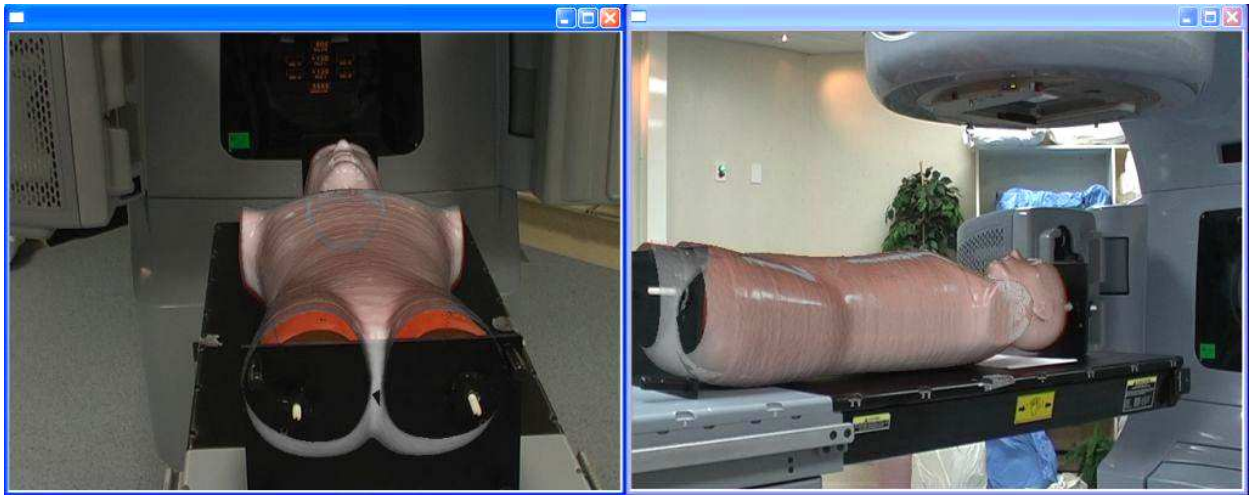
To obtain this information, a cone-beam CT scan of the phantom was obtained after set-up had been performed. This provided volumetric data of the region of interest (in this case, the phantom's head, where the treatment isocentre was located) which could be compared to the original planning CT. (The results of the comparison of the two scans are discussed in section 6.2.)



(a) The virtual phantom is displayed.



(b) The couch is raised to the correct height. Transparency and backface culling are enabled.



(c) The real phantom is shifted in the horizontal plane to align it with the virtual phantom.

Figure 6.4: The set-up process.

After the first test was performed, a second camera set-up was chosen based on the findings. An evaluation of the first set-up resolved that camera 1 (the left image in Figure 6.4) had not been elevated high enough, and the angle between camera 2 (right) and the treatment couch had been too sharp, making it difficult to gauge the phantom's position in the y-direction (towards the linac).

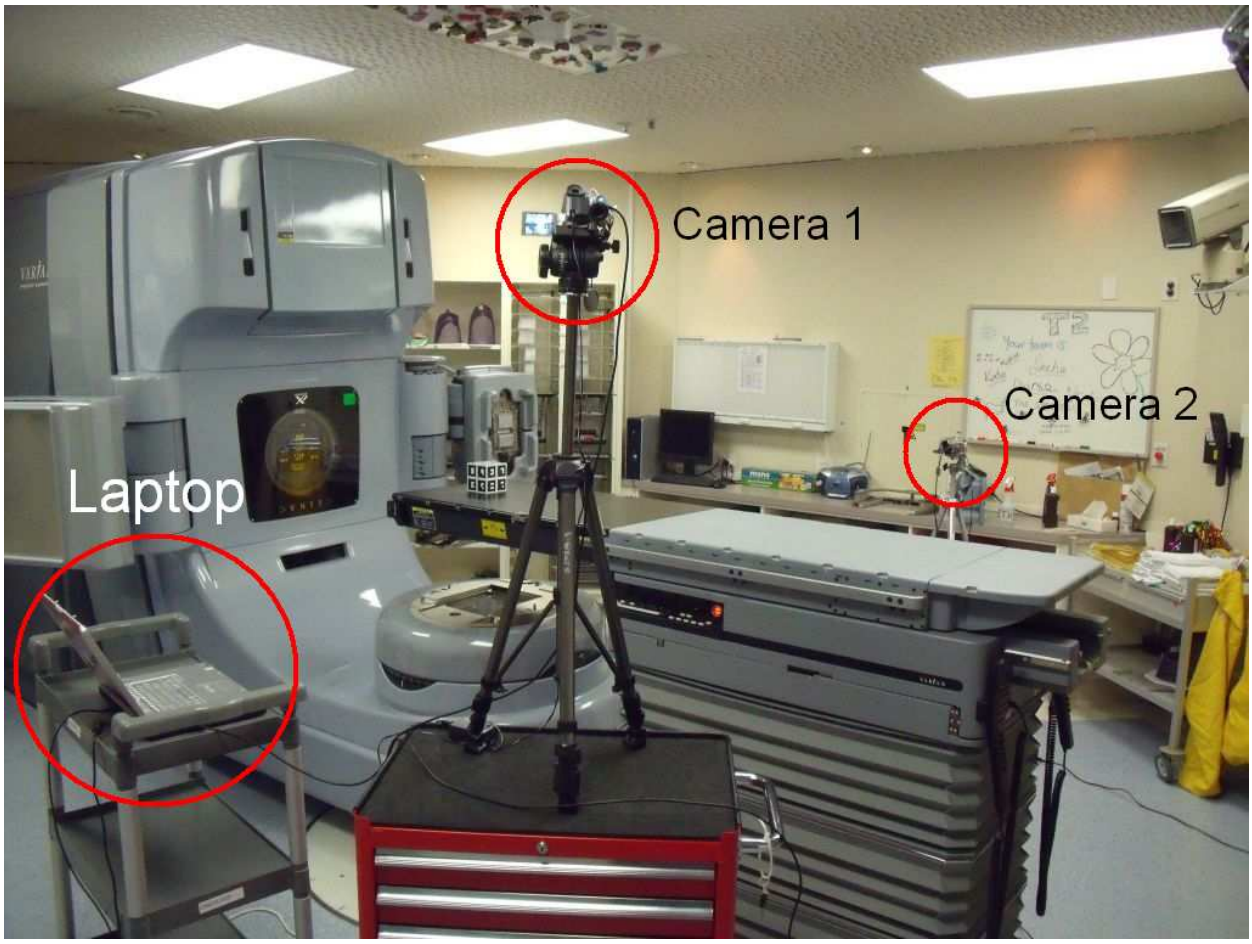
The experiment was therefore repeated with a different set-up of cameras, which is outlined in Figure 6.5. In this configuration, camera 1 was to the left of the couch, 220 cm from the isocentre and elevated 45° above it. Camera 2 was level with the isocentre, 160 cm from it, on the right-hand side of the treatment couch.

The phantom set-up process was repeated: registration was performed with the cube, the phantom was placed on the treatment couch, the couch was lowered to the correct height, and the phantom was aligned with the virtual contour (Figure 6.6). This camera configuration proved to be far more intuitive. Again, a cone-beam CT scan of the phantom was obtained to compare to the planning CT.

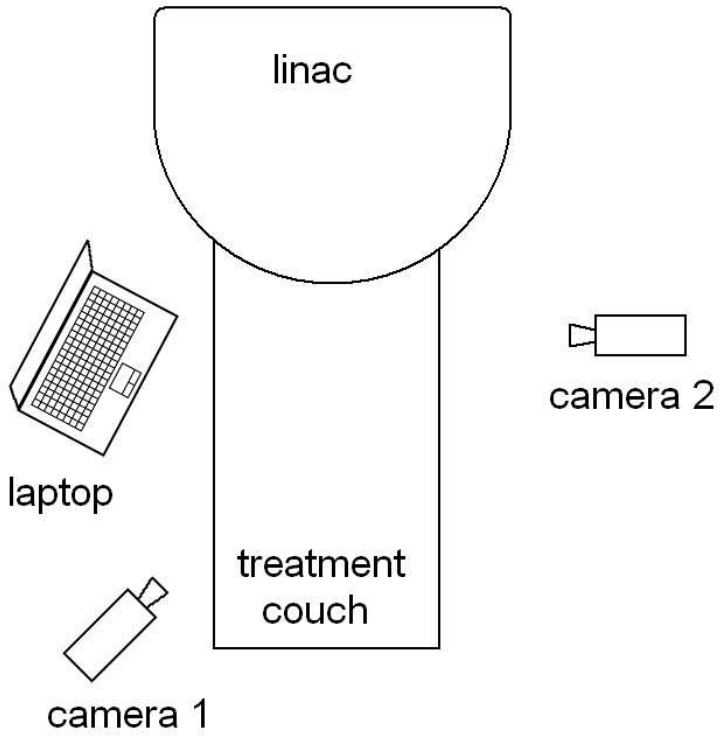
One of the intended advantages of this system was that it would provide the ability to visualise patient set-up deformations. The ability of the system to provide visualisation of non-rigid deformations was also investigated. The RANDO ® phantom had a degree of flexibility to it, and so could be slightly warped out of shape to simulate a deformation.

A head-rest was placed beneath the phantom's head, raising it a few centimetres above the surface of the treatment couch. As expected, this deformation was clearly visible in both augmented images on the laptop monitor (Figure 6.7a).

A more interesting example of non-rigid deformation is visible in Figure 6.7b. This time, the body of the phantom was contorted slightly. It was found that by shifting the treatment couch, the real phantom could still be made to apparently align with the virtual phantom in the left-hand camera image. However, in the right image, it was apparent that a set-up error had been made. This demonstrated the benefit of a two-camera system.



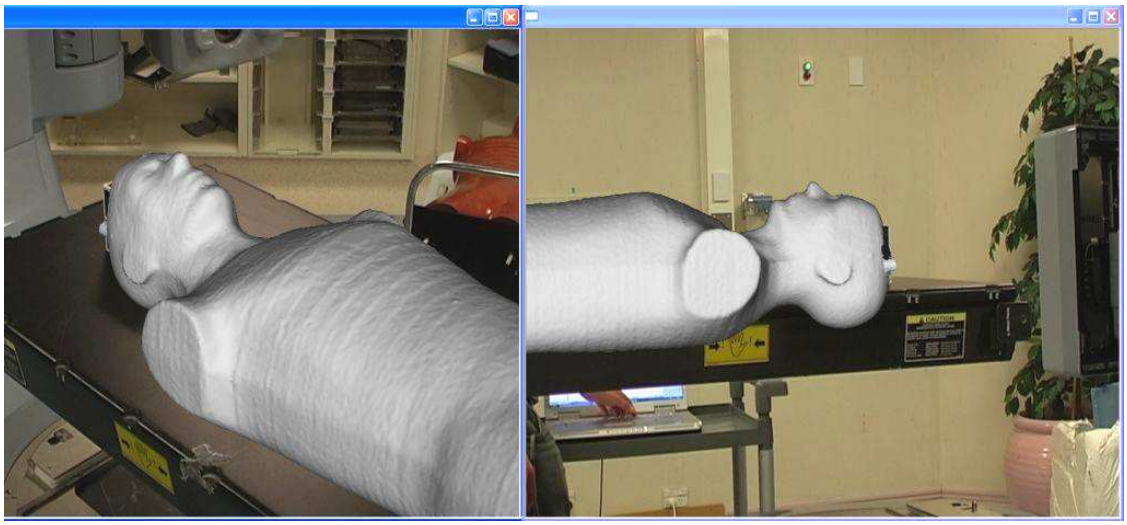
(a) Photograph of the set-up.



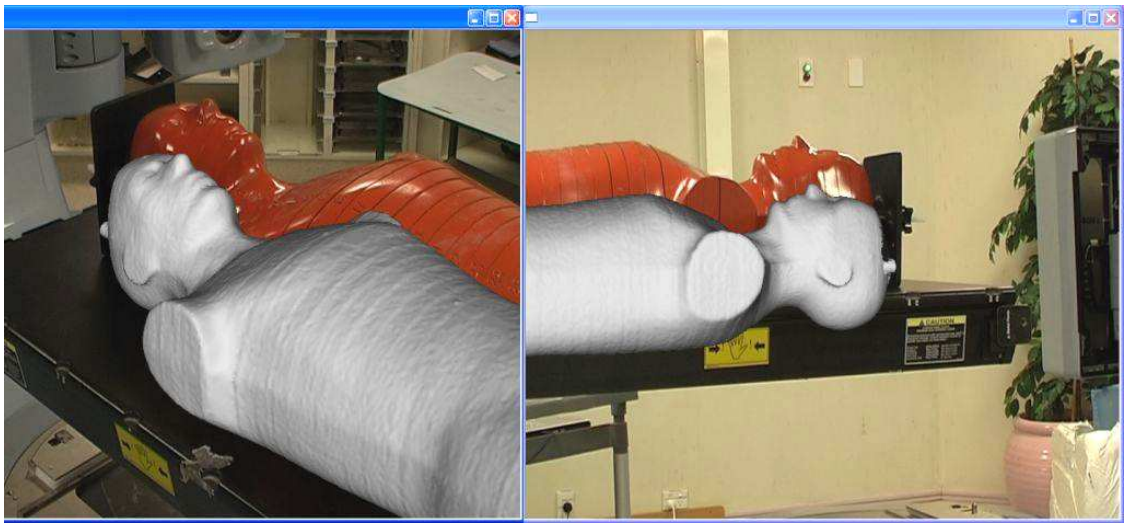
(b) Schematic of the set-up.

Figure 6.5: Second AR system set-up.

(a) The registration cube is aligned with the linac isocentre.

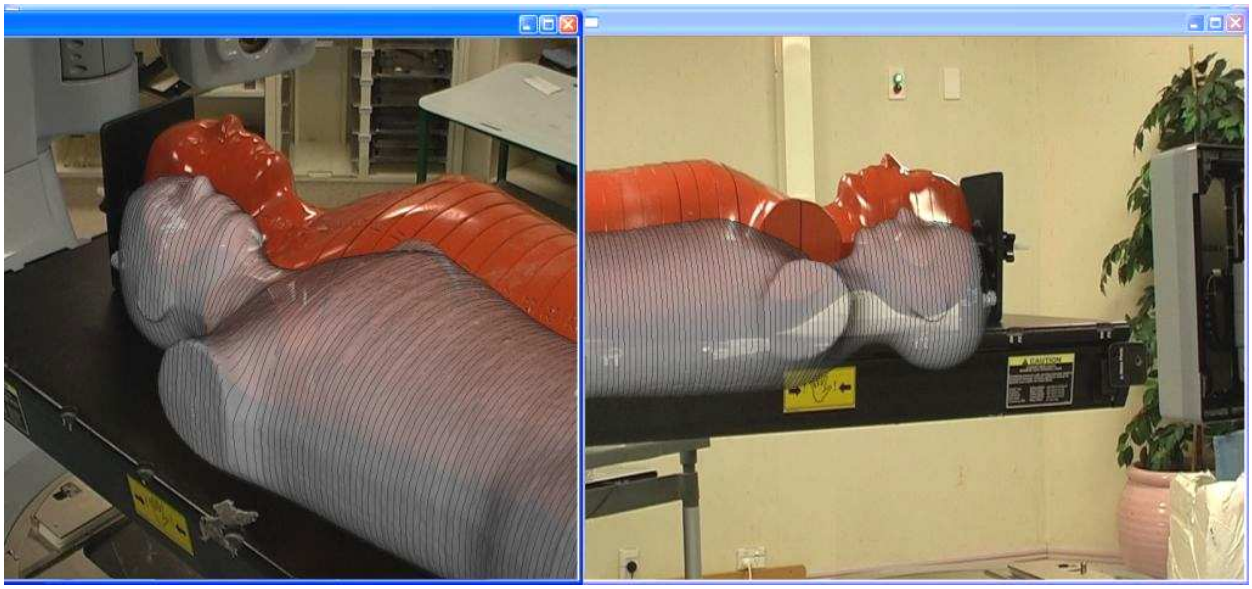


(b) The virtual phantom is displayed.

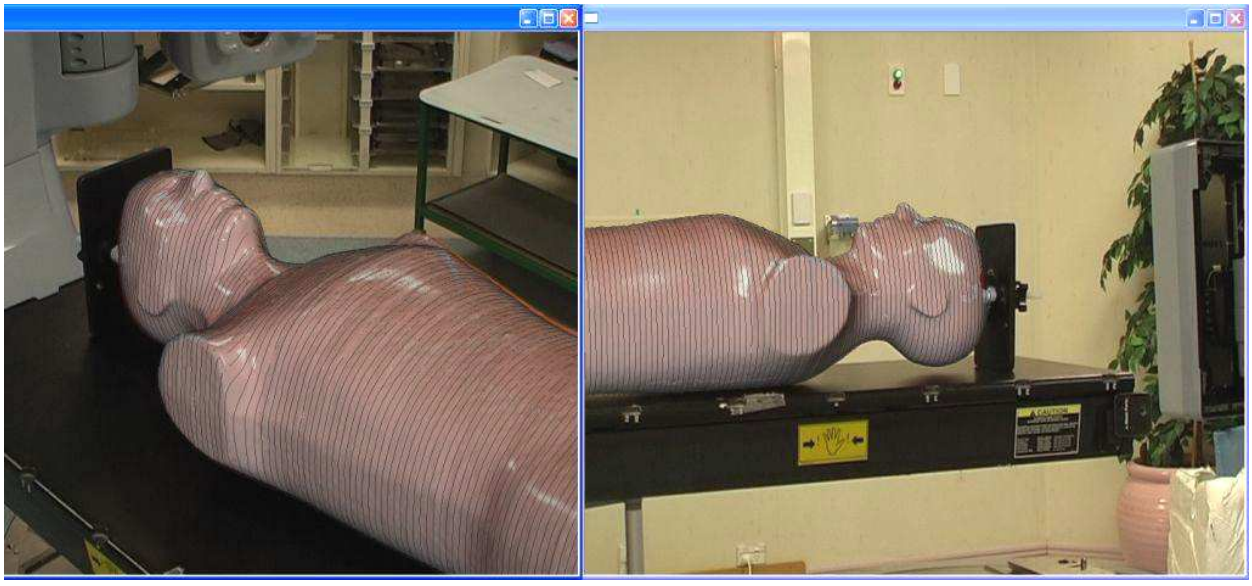


(c) The cube is removed and the real phantom is placed on the treatment couch.

Figure 6.6: Second trial of the set-up process.

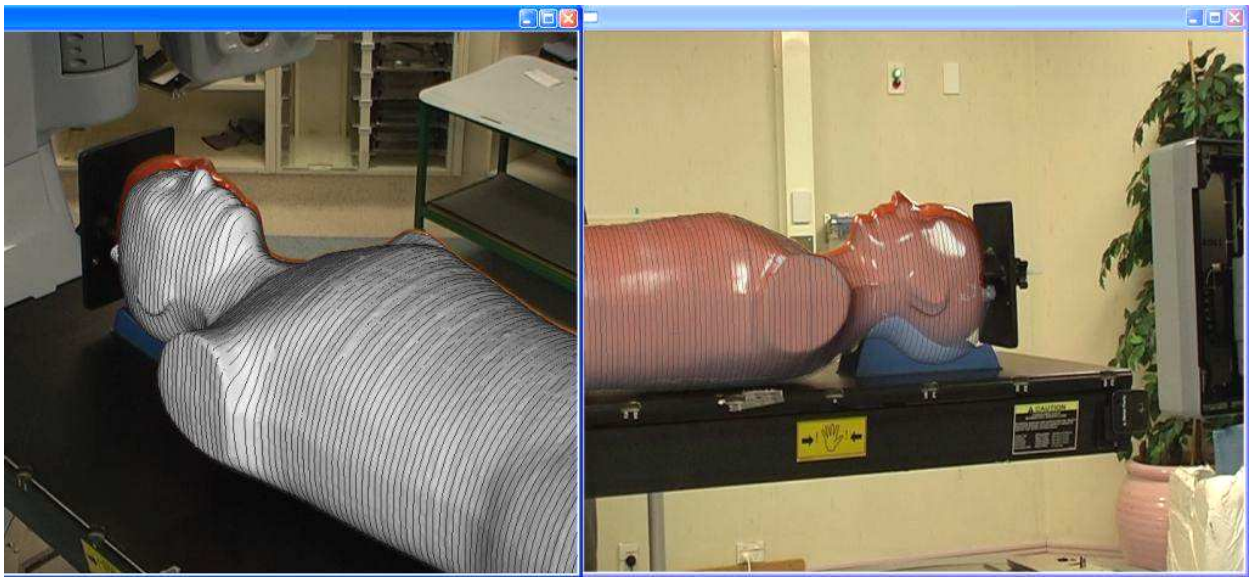


(d) 3D model transparency and wireframe are enabled to aid visualisation.

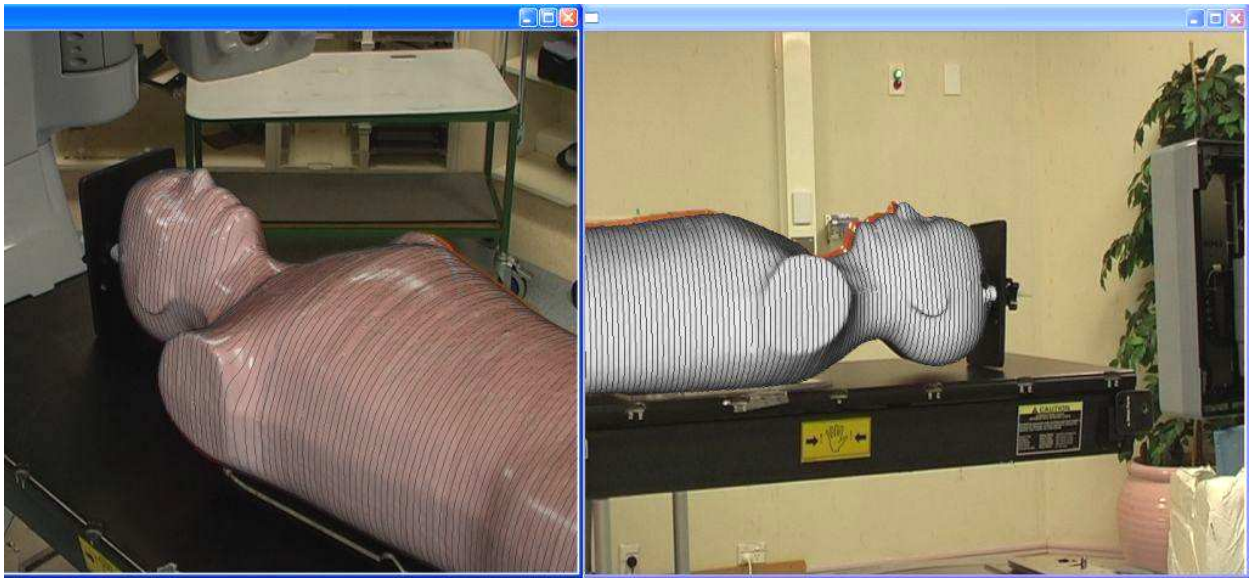


(e) The couch is lowered to the correct height and the real phantom is aligned.

Figure 6.6: (continued)



(a) The phantom's head is raised several centimetres. The discrepancy between the real and virtual phantoms is visible in the augmented images.



(b) The body of the phantom is twisted slightly. Although it seems to be perfectly aligned in the left image, there is a noticeable error in the right.

Figure 6.7: Phantom deformation.

6.2 Results

Since the RANDO ® phantom was a rigid body, its deviation from the planning CT position could be described in terms of six parameters: x, y and z, which described the linear displacement of the treatment isocentre from the linac isocentre, and pitch, roll and yaw, which described the rotation of the phantom on the treatment couch (Figure 6.8). Note that these are the axes of the CT scan data, which are oriented differently to those that were defined for the spatial registration process in section 2.4.

To determine the accuracy of the set-up with the AR system, these parameters were required. To obtain them, a software package for MATLAB called "SPM" (Statistical Parametric Mapping) was used [62]. Although SPM is intended for the analysis of brain imaging

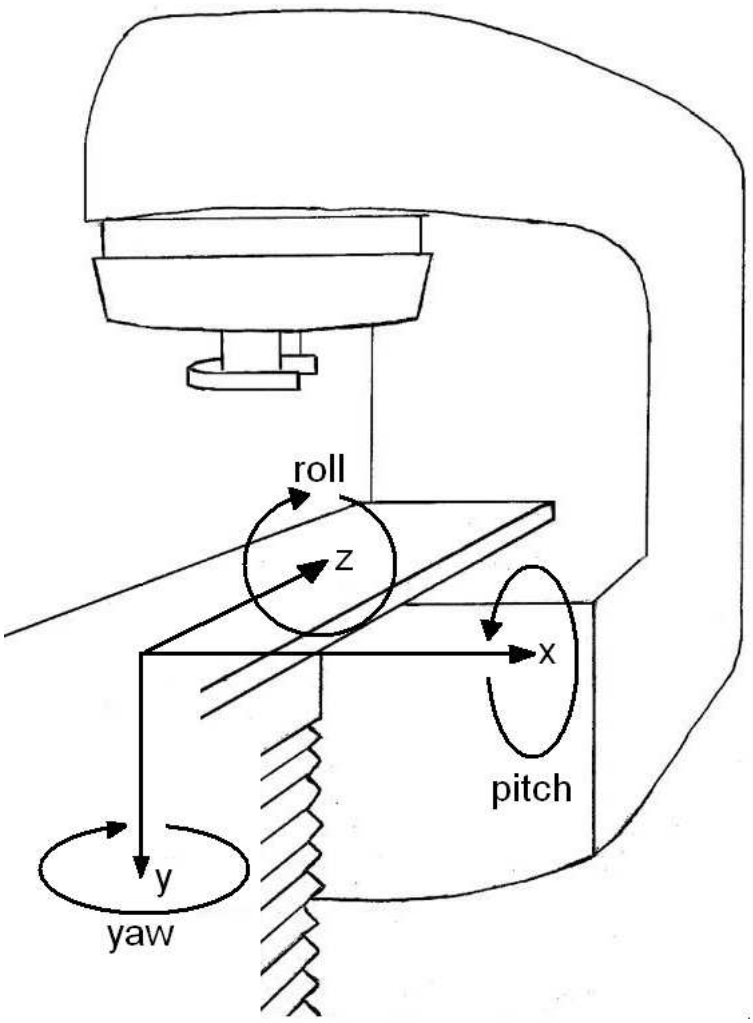


Figure 6.8: The six parameters that describe the position of a rigid body on the treatment couch.

data sequences, it provides a function that allows volumetric datasets to be aligned with one another. It uses a least-squares algorithm and a rigid-body spatial transformation, and outputs the six parameters associated with the realignment, and so was ideal for obtaining the results of this experiment.

The origin of the coordinates in the planning CT was at the planning isocentre. In the CBCT scans, the origin was set to the linac isocentre. As a result of the set-up error, when the CT and CBCT volumes were registered together so that their origins coincided, they were slightly offset.

The SPM realignment algorithm was applied to the datasets to bring them into spatial alignment (as shown in Figure 6.9), and the set-up error parameters were obtained. The results for both camera arrangements are listed in Table 6.1. The magnitudes of the translational errors were obtained using Pythagoras' theorem, as follows:

$$magnitude = \sqrt{x^2 + y^2 + z^2}$$

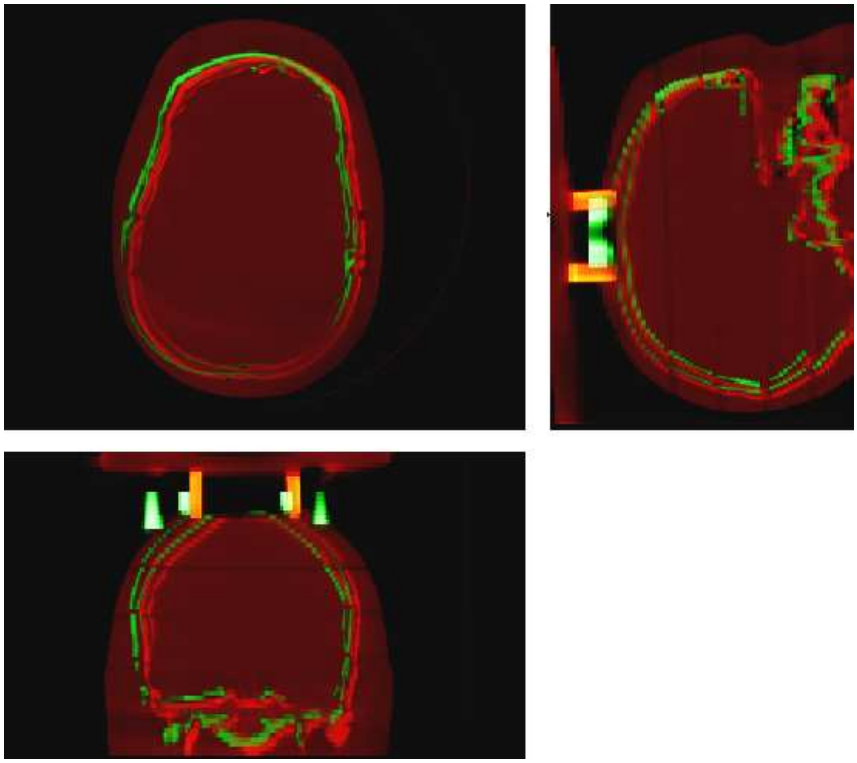
$$magnitude_1 = \sqrt{(-4.3)^2 + 1.8^2 + (-3.6)^2} \qquad magnitude_2 = \sqrt{(-2.4)^2 + (-1.7)^2 + 0.7^2}$$

$$magnitude_1 = 5.9mm \qquad magnitude_2 = 3.0mm$$

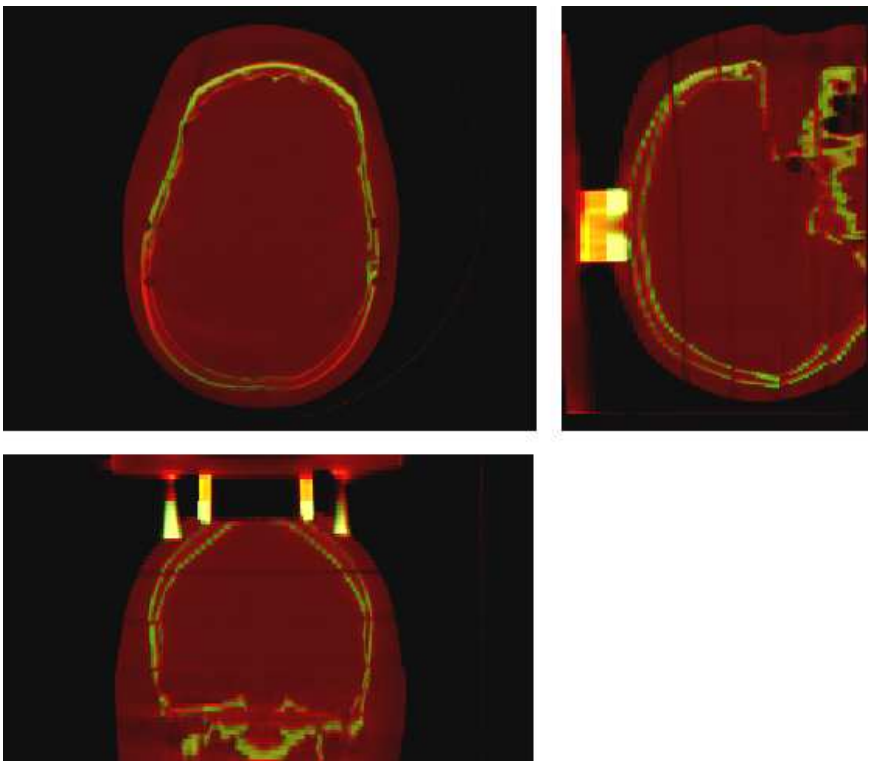
As expected, the error from the second set-up was smaller, owing to the optimised camera arrangement. Interestingly, in both camera arrangements, the translational error was greatest in the x direction. A possible cause for this was a systematic error in the tracking software, such as an undetected error in the calibration parameters. This could have caused the 3D model representation to be displayed slightly, but not discernibly, offset in the x direction.

Set-up	x (mm)	y (mm)	z (mm)	Error magnitude (mm)	pitch (°)	roll (°)	yaw (°)
1	-4.3	1.8	-3.6	5.9	-0.22	-0.23	0.11
2	-2.4	-1.7	0.7	3.0	-0.19	-0.06	0.27

Table 6.1: Translational and rotational offset of phantom set-up.



(a) The scan volumes prior to realignment. There is a noticeable discrepancy between the positions of bony structures.



(b) After realignment.

Figure 6.9: Alignment of the CT (green) and CBCT (red) volumes. The transverse plane is shown in the top-left images, the sagittal in the top-right, and the coronal in the bottom-left. Both views were obtained using AMIDE [63], an open-source program for the examination and comparison of medical imaging data.

Chapter 5.4 mentioned that the uncertainty in the room-laser intersection point was 2 mm. This was another possibility for the source of the systematic error. A study by Meyer *et al* [64] reported that positioning accuracy with cone-beam CT could be performed with translational errors as low as 0.2 mm, so if the cube had been aligned using CBCT instead of just room lasers, it is possible that the translational error could have been greatly reduced.

The results of the experiment were acceptable, but there was no guarantee that repetition of the experiment would produce similar results. In other words, it was repeated too few times to ensure that the same degree of accuracy could be reproduced indefinitely, and that the results were irrefutably representative of the system. Repetition of the experiment, preferably with a range of users, would provide a more confident estimate of the average set-up uncertainty that could be obtained.

Nevertheless, given that the system was in its early stages of development during this experiment, the 3 mm translational error was well within acceptable limits. Similarly, the rotational set-up error was small, indicating that the system performed pose-estimation of the 3D model with satisfactory accuracy.

Chapter 7

Discussion and Conclusion

This project has shown that augmented reality is entirely viable as a tool for patient positioning in radiotherapy. A method of coordinate registration was developed using the registration cube and fiducial tracking markers, and a range of visualisation techniques were used to allow intuitive visual guidance. Experiments were performed in a radiotherapy clinic, which validated the system's performance and provided satisfying accuracy results.

Throughout development, a number of minor limitations of the system became apparent. These issues must be addressed in future work before it becomes ready to be fully implemented in a clinical environment.

The operation of the system as a whole was separated into several steps, involving the use of a number of computer programs. The user first obtains the pointset from the planning CT using a MATLAB script, then performs surface reconstruction using Meshlab. Camera calibration is performed using an application included with the ARToolKit tracking library, and finally the AR application itself is executed. All of these processes should be integrated into a single application, with an intuitive graphical user interface (GUI).

The experiment that was described in the previous chapter evaluated the accuracy with which the system allowed patient set-up to be performed. However, it was carried out by the developers, who were already familiar with its operation and knew how to optimise its performance. To obtain an objective appraisal of the system, a regime of usability tests should be devised. This would involve the participation of several staff members at a radiotherapy

facility to try the system.

The accuracy of the system will not be the only factor to consider if usability tests are carried out. The system was intended to allow the detection of non-rigid deformations to patient pose, so that they could be corrected before rigid-body corrections are applied with IGRT. This ability has not been investigated in great depth, so feedback on this particular issue should be sought.

The system was also designed to allow variations to patient geometry to be observed. Over the course of fractionated treatment (typically six weeks), changes to patient geometry occur due to factors such as weight loss or gain, or swelling. There will also be changes within the timeframe of a single treatment due to respiration, relaxing of muscles, and movement due to discomfort. There will therefore be some discrepancy between the patient's external contour in the planning CT and during treatment. The ability to observe the magnitude of these changes should be investigated, so the usability tests should include extensive testing with real patients.

The usability tests should also seek feedback on more subjective issues, specifically the interface of the system. The success of an AR application depends on the seamlessness of human interaction. For this reason, efforts should be made to improve the representation of the 3D model in the augmented image. Chapter 3 discussed the techniques that were investigated in this project, but to allow further development in this area, user experience data must be obtained. This would indicate the shortcomings of the interface, and measures could be taken to improve it. Since it is difficult to measure intuition for evaluation quantitatively, a survey of subjective questions relating to the user's experience should be prepared.

The system prototype is still in early days of development, and a range of suggestions have been made for future improvements to the system, so that its potential can be fully realised.

Perhaps the most significant issue that was identified was the accuracy of the registration process. The "registration cube" was a novel idea which allowed the AR software to directly track the linac isocentre. However, the tracking accuracy was hindered by the small size

of the cube, and the tracking markers attached to its faces. When displaying larger (i.e. patient sized) objects, the tracking software had to extrapolate further in the geometric transformation calculations, resulting in translational and rotational registration errors (as discussed in section 5.2). The simple solution to this problem would be to construct a larger cube, with larger tracking markers. As an estimate, a cube with 30 cm dimensions should suffice.

If a cube of this size were constructed, one must consider the practical implications. A larger cube would definitely allow more accurate registration, but would be heavier and more cumbersome, and possibly even difficult to lift onto the treatment couch. Alternatively, the cube could be made hollow, or of a lighter material than perspex, but the possibility of warping over time must be kept in mind.

As noted in section 5.3, the cameras displayed a prominent distortion effect that calibration was not able to entirely compensate for. This placed another limitation on the accuracy of the system. Although care was taken to position the registration cube at the centre of the camera images during the hospital tests, the distortion effect would have had an impact on the phantom set-up, and may have been a contributing factor to the angular set-up error.

A type of camera that suffers from a minimal amount of distortion would be ideal for this system. The camera resolution should also be taken into account, since a sharper camera image allows the tracking software to more precisely locate the tracking markers.

As an alternative to bringing new cameras into the linac room, there is the possibility of accessing the cameras that are already in the linac bunker. Most bunkers contain two of these cameras, as a way of monitoring the patient throughout treatment. However, the use of these would require that they remain stationary. The zoom levels would also have to remain unaltered, since zoom affects camera calibration. This could be considered motivation to develop a method of tracking after the registration cube is removed (perhaps by means of markerless AR tracking [65]) and to compensate for zoom dynamically, without having to recalibrate the camera whenever it is altered.

The nature of this project has been largely investigative, and in this regard in particular it has been a success. A number of shortcomings with the system have been identified and

discussed, as well as suggestions for improvements that could be made in future projects. Overall, the project has proven the feasibility of augmented reality as a device for patient position guidance in radiotherapy.

Experiments with the system have produced acceptable results, and have shown that it can be used to perform set-up accurately. This has in turn shown that, in principle, the methodology that has been employed to develop the system has been sound. If the appropriate measures are taken to build upon the findings of this project in future works, it could be made into an indispensable tool for radiotherapy procedures.

Bibliography

[1] World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. *The World Health Organization's Fight Against Cancer: Strategies That Prevent, Cure and Care*, 2007.

[2] Cancer Control Taskforce. *The New Zealand Cancer Control Strategy: Action Plan 2005–2010*. Wellington: Ministry of Health, PO Box 5013, Wellington, New Zealand, 2005.

[3] Wellington: Ministry of Health, PO Box 5013, Wellington, New Zealand. *Improving Non-Surgical Cancer Treatment Services in New Zealand*, 2001.

[4] G. Delaney, S. Jacob, C. Featherstone, and M. Barton. The Role of Radiotherapy in Cancer Treatment: Estimating Optimal Utilization from a Review of Evidence-Based Clinical Guidelines. *Cancer*, 104(6):1129–1137, 2005.

[5] M. K. Bucci, A. Bevan, and M. Roach III. Advances in Radiation Therapy: Conventional to 3D, to IMRT, to 4D, and Beyond. *Ca- A Cancer Journal for Clinicians*, 55(8):117–134, 2005.

[6] E. B. Podgorsak. *Radiation Oncology Physics: A Handbook for Teachers and Students*. International Atomic Energy Agency, 2005.

[7] A. L. Boyer. Laser "cross-hair" sidelight. *Medical Physics*, 5(1):58–60, 1978.

[8] I. A. Brezovich and S. Jordan. A device for precision positioning and alignment of room lasers to diminish their contribution to patient setup errors. *Journal of Applied Clinical Medical Physics*, 8(4):45–53, 2007.

[9] C. W. Hurkmans, P. Remeijer, J. V. Lebesque, and B. J. Mijnheer. Set-up verification using portal imaging; review of current clinical pracitce. *Radiotherapy and Oncology*, 58(2):105–120, 2001.

[10] E. J. Hall and A. J. Giaccia. *Radiobiology for the Radiobiologist*. Lippincott Williams & Wilkins, 2006.

[11] M. Elshaikh, M. Ljungman, R. T. Haken, and A. S. Lichter. Advances in Radiation Oncology. *Annual Review of Medicine*, 57(1):19–31, 2006.

[12] G. Budgell. Intensity modulated radiotherapy (IMRT)—an introduction. *Radiography*, 8:241–249, 2002.

[13] L. A. Dawson and M. B. Sharp. Image-guided radiotherapy: rationale, benefits, and limitations. *The Lancet Oncology*, 7(10):848–858, 2006.

[14] D. Verellen, M. De Ridder, K. Tournel, M. Duchateau, T. Reynders, T. Gevaert, N. Linthout, and G. Storme. An overview of volumetric imaging technologies and their quality assurance for IGRT. *Acta Oncologica*, 47(7):1271–1278, 2008.

[15] M. Guckenberger, J. Meyer, D. Vordermark, K. Baier, J. Wilbert, and M. Flentje. Magnitude and Clinical Relevance of Translational and Rotational Patient Setup Errors: A Cone-beam CT Study. *International Journal of Radiation Oncology Biology Physics*, 65(3):934–942, 2006.

[16] K. K. Brock, M. Hawkins, C. Eccles, J. L. Moseley, D. J. Moseley, D. A. Jaffray, and L. A. Dawson. Improving image-guided localization through deformable registration. *Acta Oncologica*, 47(7):1279–1285, 2008.

[17] B. Fei, C. Krempien, and D. L. Wilson. A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes. *Computerized Medical Imaging and Graphics*, 27(4):267–281, 2003.

[18] B. Fei, A. Wheaton, Z. Lee, J. L. Duerk, and D. L. Wilson. Automatic MR volume registration and its evaluation for the pelvis and prostate. *Physics in Medicine and Biology*, 47(5):823–838, 2002.

[19] M. Van Herk, A. Bruce, A. P. G. Kroes, T. Shouman, A. Touw, and J. V. Lebesque. Quantification of organ motion during conformal radiotherapy of the prostate by three

dimensional image registration. *International Journal of Radiation Oncology Biology Physics*, 33(5):1311–1320, 1995.

[20] S. Chopra, K. A. Dinshaw, R. Kamble, and R. Sarin. Breast movement during normal and deep breathing, respiratory training and set up errors: implications for external beam partial breast irradiation. *The British Journal of Radiology*, 79:766–773, 2006.

[21] E. J. Harris, E. M. Donovan, J. R. Yarnold, C. E. Coles, P. M. Evans, and IMPORT Trial Management Group. Characterisation of Target Volume Changes During Breast Radiotherapy Using Implanted Fiducial Markers and Portal Imaging. *International Journal of Radiation Oncology Biology Physics*, 73(3):958–966, 2009.

[22] B. Polat, J. Wilbert, K. Baier, M. Flentje, and M. Guckenberger. Nonrigid Patient Setup Errors in the Head-and-Neck Region. *Strahlentherapie und Onkologie*, 183:506–511, 2007.

[23] R. T. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[24] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.

[25] R. Behringer, J. Christian, A. Holzinger, and S. Wilkinson. Some Usability Issues of Augmented and Mixed Reality for e-Health Application in the Medical Domain. *Lecture Notes in Computer Science*, 4799:255–266, 2007.

[26] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced Medical Displays: A Literature Review of Augmented Reality. *IEEE/OSA Journal of Display Technology*, 4(4):451–467, 2008.

[27] M. Harders, G. Bianchi, and B. Knoerlein. Multimodal Augmented Reality in Medicine. *Lecture Notes in Computer Science*, 4555:652–658, 2007.

[28] D. G. Trevisan, L. P. Nedel, and B. Macq. Augmented Vision for Medical Applications. *Proceedings of the ACM Symposium on Applied Computing*, pages 1415–1419, 2008.

[29] H. Deutschmanm, P. Steininger, O. Nairz, P. Kopp, F. Merz, K. Wurstbauer, F. Zehentmayr, G. Fastner, M. Krazinger, G. Kametriser, M. Kopp, and F. Sedlmayer.

"Augmented Reality" in Conventional Simulation by Projection of 3-D Structures into 2-D Images: A Comparison with Virtual Methods. *Strahlentherapie und Onkologie*, 184(2):93–99, 2008.

[30] R. Krempien, H. Hoppe, L. Kahrs, S. Daeuber, O. Schorr, G. Eggers, M. Bischof, M. W. Munter, J. Debus, and W. Harms. Projector-Based Augmented Reality for Intuitive Intraoperative Guidance in Image-Guided 3D Interstitial Brachytherapy. *International Journal of Radiation Oncology Biology Physics*, 70(3):944–952, 2008.

[31] A. P. Santhanam, T. R. Willoughby, I. Kaya, A. P. Shah, S. L. Meeks, J. P. Rolland, and P. A. Kupelian. A Display Framework for Visualizing Real-Time 3D Lung Tumour Radiotherapy. *IEEE/OSA Journal of Display Technology*, 4(4):473–482, 2008.

[32] S. Wesarg, E. A. Ferle, B. Schwald, H. Seibert, P. Zogal, and S. Roeddiger. Accuracy of needle implantation in brachytherapy using a medical AR system - a phantom study. *Proceedings of SPIE International Symposium on Medical Imaging*, 5367:341–352, 2004.

[33] A. Brahme, P. Nyman, and B. Skatt. 4D laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures. *Medical Physics*, 35(5):1670–1681, 2008.

[34] C. Bert, K. G. Metheany, K. Doppke, and G. T. Y. Chen. A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Medical Physics*, 32(9):2753–2762, 2005.

[35] M. F. Spadea, G. Baroni, M. Riboldi, B. Tagaste, C. Garibaldi, R. Orecchia, and A. Pedotti. Patient set-up verification by infrared optical localization and body surface sensing in breast radiation therapy. *Radiotherapy and Oncology*, 79:170–178, 2006.

[36] Jr W. D. Bidgood, S. C. Horii, F. W. Prior, and D. E. Van Syckle. Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging. *Journal of the American Medical Informatics Association*, 4(3):199–212, 1997.

[37] The Mathworks Inc. MATLAB version 7.5.0.342 (R2007b). <http://www.mathworks.com/products/matlab/>, accessed February 2009, 2007.

[38] The Mathworks Inc. Image Processing Toolbox version 6.0 (R2007b). <http://www.mathworks.com/products/image/>, accessed February 2009, 2007.

[39] P. Bourke. Object Files (.Obj). <http://local.wasp.uwa.edu.au/ pbourke/dataformats/obj/>, accessed February 2009.

[40] P. Bourke. Minimum Requirements For Creating A DXF File Of A 3D Model. <http://local.wasp.uwa.edu.au/ pbourke/dataformats/dxf/min3d.html>, accessed February 2009, 1993.

[41] P. Cignoni, M. Callieri, M. Corsini, F. Ganovelli, and G. Ranzuglia. Meshlab v1.1.1, Visual Computing Lab - ISTI - CNR. http://meshlab.sourceforge.net/, march 2008.

[42] F. Bernadini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin. The Ball-Pivoting Algorithm for Surface Reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 5(4):349–359, 1999.

[43] J. Looser, R. Grasset, H. Seichter, and M. Billinghurst. OSGART - A Pragmatic Approach to MR. *Industrial Workshop at ISMAR 2006, Santa Barbara, California*, 2006.

[44] OpenSceneGraph. <http://www.openscenegraph.org/projects/osg>, accessed February 2009.

[45] ARToolworks Inc, Human Interface Technology Laboratory (HITLab), and Human Interface Technology Laboratory New Zealand (HITLabNZ). ARToolKit Professional version 4.3. <http://www.artoolworks.com/Home.html>, accessed February 2009.

[46] H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, pages 85–94, 1999.

[47] O. Bimber and R. Raskar. Alternative Augmented Reality Approaches: Concepts, Techniques, and Applications. Eurographics 2003, <http://www.uni-weimar.de/medien/ar/SpatialAR/EG03.htm>, accessed February 2009.

[48] Y. Yokokohji, Y. Sugawara, and T. Yoshikawa. Accurate Image Overlay on Video See-Through HMDs Using Vision and Accelerometers. *Proceedings of the Virtual Reality Annual International Symposium*, pages 247–254, 2000.

[49] E. McGarrity and M. Tuceryan. A Method for Calibrating See-through Head-mounted Displays for AR. *Proceedings of the IEEE 2nd International Workshop on Augmented Reality (IWAR 99)*, pages 75–84, 1999.

[50] Brown University. Methods For Improving Outcomes in Surgical Procedures, 2008. <http://biomed.brown.edu/Courses/BI108/BI108_2008_Groups/group12/surgicaltraining.html>, accessed February 2009.

[51] G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. S. Fischer, H. Mathieu, R. H. Taylor, J. Zinreich, and L. M. Fayad. Image Overlay Guidance for Needle Insertion in CT Scanner. *IEEE Transactions on Biomedical Engineering*, 52(8):1415–1424, 2005.

[52] 3D Visualisation and Imaging System Lab. homepage. <http://3dvis.optics.arizona.edu>, accessed February 2009.

[53] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):102–107, 1998.

[54] H. Kato, Human Interface Technology Laboratory (HITLab), Human Interface Technology Laboratory New Zealand (HITLabNZ), and ARToolworks Inc. ARToolKit version 2.65. <http://www.hitl.washington.edu/artoolkit>, accessed February 2009.

[55] Open Source Computer Vision (OpenCV) wiki. <http://opencv.willowgarage.com/wiki/>, accessed February 2009.

[56] M. Tuceryan, D. S. Greer, R. T. Whitaker, D. E. Breen, C. Crampton, E. Rose, and K. H. Ahlers. Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System. *IEEE Transactions on Visualisation and Computer Graphics*, 1(3):255–273, 1995.

[57] D. Stoyanov, G. P. Mylonas, M. Lerotic, A. J. Chung, and G.-Z. Yang. Intra-Operative Visualizations: Perceptual Fidelity and Human Factors. *Journal of Display Technology*, 4(4):491–501, 2008.

[58] J. O. Deasy and Washington University in St. Louis School of Medicine. CERR: Computational Environment for Radiotherapy Research, version 3.0 beta 5, 2007. <http://radium.wustl.edu/CERR/about.php>, accessed February 2009.

[59] Y. Nishina, B. Okumura, M. Kanbara, and N. Yokoya. Photometric registration by adaptive high dynamic range image generation for augmented reality. *IEEE International Symposium on Mixed and Augmented Reality*, pages 53–56, 2008.

[60] G. J. Kutcher, L. Coia, M. Gillin, W. F. Hanson, S. Leibel, R. J. Morton, J. R. Palta, J. A. Purdy, L. E. Reinstein, G. K. Svensson, M. Weller, and L. Wingfield. Comprehensive QA for Radiation Oncology: Report of AAPM Radiation Therapy Committee Task Group 40. *Medical Physics*, 21(4):581–618, 1994.

[61] The Phantom Laboratory. *Technical data sheet for the RANDO ® phantom.* <http://www.phantomlab.com/rando.html>, accessed February 2009.

[62] The Wellcome Trust Centre for Neuroimaging. SPM5. <http://www.fil.ion.ucl.ac.uk/spm/>, accessed February 2009.

[63] A. Loening. AMIDE, (A Medical Image Data Examiner), version 0.9,1, 2007. <http://amide.sourceforge.net>, accessed February 2009.

[64] J. Meyer, J. Wilbert, K. Baier, M. Guckenberger, A. Richter, O. Sauer, and M. Flentje. Positioning accuracy of cone-beam computed tomography in combination with a hexapod robot treatment table. *International Journal of Radiation Oncology Biology Physics*, 67(4):1220 – 1228, 2007.

[65] C. Yuan. Markerless pose tracking for augmented reality. *Lecture Notes in Computer Science*, 4291:721 – 730, 2006.